

Reports in Phonetics, University of the Saarland  
Berichte zur Phonetik, Universität des Saarlandes

## PHONUS

Edited by: W. J. Barry, B. Möbius & J. Trouvain

No. 18, March 2015

**Institut für Phonetik  
Universität des Saarlandes  
D-66041 Saarbrücken, März 2015**

**ISSN 0949-1791**



## VORWORT DER HERAUSGEBER

PHONUS 18 setzt die Reihe der Veröffentlichungen von Dissertationen von Mitgliedern der Saarbrücker Phonetik-Gruppe fort. Fabian Brackhane, wissenschaftlicher Mitarbeiter des Instituts für Deutsche Sprache in Mannheim und Doktorand an der Universität des Saarlandes, legt hier seine Doktorarbeit mit dem Titel »Kann was natürlicher, als Vox humana, klingen?« Ein Beitrag zur Geschichte der mechanischen Sprachsynthese« vor.

Fabian Brackhane befasst sich mit der Entwicklungsgeschichte der mechanischen Sprachsynthese ab dem 17. Jahrhundert, unter besonderer Berücksichtigung der in eine wissenschaftliche Sackgasse führenden Synthesekonzepte auf der Grundlage von Prinzipien des Orgelbaus. Wie Brackhane zeigt, hielt sich die irrige Vorstellung, in der Orgel und hier insbesondere im Orgelregister Vox Humana ein geeignetes Vorbild für die Sprachsynthese zu besitzen, hartnäckig und mindestens bis ins 19. Jahrhundert. Der Zusammenhang zwischen Sprachsynthese und Orgelbau wurde auch in der Preisfrage der Russischen Akademie der Wissenschaften über die Natur der menschlichen Vokale nicht nur betont, sondern sogar eingefordert. Mindestens vier Forscher arbeiteten gegen Ende des 18. Jahrhunderts zeitgleich und weitgehend unabhängig voneinander an Sprachmaschinen, die das Vox-Humana-Konzept aufgriffen. Und selbst umfassend gebildete Wissenschaftler wie Kempelen, der Erfinder der - aus heutiger Sicht erfolgreichsten - Sprachmaschine mit dem größten Grad der Analogie zu den menschlichen Sprechorganen, erlag noch dem Mythos der Vox Humana.

Diese Dissertation stellt erstmals die technologischen Konzepte der mechanischen Sprachsynthese in einen systematischen Zusammenhang mit ihren Vorbildern im Orgelbau. Aufschlussreich sind auch die experimentalphonetischen Untersuchungen der akustischen Struktur von Orgeltonen, die mit der Pfeifenbauform Vox Humana erzeugt wurden, und ihrer - wie sich zeigt: doch eher vagen - akustischen und perzeptuellen Ähnlichkeit zu menschlichen Sprachlauten Vokalen.

Saarbrücken, im März 2015     William J. Barry, Bernd Möbius & Jürgen Trouvain



# »Kann was natürlicher, als Vox humana, klingen?«

Ein Beitrag zur Geschichte der  
mechanischen Sprachsynthese

Dissertation  
zur Erlangung des akademischen Grades eines  
Doktors der Philosophie  
der Philosophischen Fakultäten  
der Universität des Saarlandes

vorgelegt von

Fabian Brackhane

aus Regensburg



Dekan: Prof. Dr. Ralf Bogner

Berichterstatter: Prof. Dr. Bernd Möbius  
Prof. em. Dr. William J. Barry  
Prof. em. Dr.-Ing. Rüdiger Hoffmann

Tag der letzten Prüfungsleistung: 21.01.2015

### **Short summary**

This dissertation deals with the development of mechanical speech synthesis from its speculative origins to the 20th century, focusing on relevant advances in the 18th century, which continued to be influential for a long time. In addition to the outstanding representatives of that time – WOLFGANG VON KEMPELEN and CHRISTIAN GOTTLIEB KRATZENSTEIN – the work also discusses a number of lesser-known 18th and 19th century speech synthesis researchers. The actual historical relationship between the mechanical speech synthesis and the organ stop *Vox humana* is presented in detail.

In addition to this extensive historical treatise, acoustic analyses and perception tests are presented which have been devised to investigate the often suggested similarity of the *Vox Humana* to the human voice. The work concludes with a set of partially already implemented experimental studies on the synthesis concepts of KEMPELEN and KRATZENSTEIN.

### **Kurzzusammenfassung**

Diese Arbeit befasst sich mit der Entwicklung der mechanischen Sprachsynthese von ihren legendenhaft-spekulativen Anfängen bis ins 20. Jahrhundert. Der Schwerpunkt der Darstellung liegt hierbei auf den für lange Zeit maßgeblichen Entwicklungen des 18. Jahrhunderts. Neben den herausragenden Vertretern dieser Zeit – WOLFGANG VON KEMPELEN und CHRISTIAN GOTTLIEB KRATZENSTEIN – wird auch auf die zahlreichen weniger bekannten Sprachsyntheseforscher des 18. und 19. Jahrhunderts eingegangen. Der faktische historische Zusammenhang zwischen der mechanischen Sprachsynthese und dem Orgelregister *Vox humana* wird eingehend dargestellt.

Neben dieser umfangreichen historischen Abhandlung werden akustische Analysen und Perzeptionsexperimente vorgestellt, mit denen die oftmals suggerierte Ähnlichkeit der *Vox humana* mit der menschlichen Stimme untersucht wurde. Bereits durchgeführte und projektierte experimentelle Untersuchungen zu den Synthese-Konzepten KEMPELENS und KRATZENSTEINS bilden den Abschluss.



# Inhaltsverzeichnis

<b>Einleitung</b>	<b>XIII</b>
<b>I Mechanische Sprachsynthese – ein historischer Überblick</b>	<b>1</b>
<b>1 Frühe Berichte zu Sprachsynthesen</b>	<b>3</b>
<b>2 Das Orgelregister »Vox humana«</b>	<b>9</b>
2.1 Etymologie und Geschichte . . . . .	9
2.2 Allgemeine Konstruktionsmerkmale . . . . .	19
2.3 Ein Prototyp der Sprachsynthese? . . . . .	22
2.4 Zusammenfassung . . . . .	27
<b>3 18. Jh.: Grundlegende Konzepte</b>	<b>29</b>
3.1 Erasmus Darwin . . . . .	29
3.2 Christian Gottlieb Kratzenstein . . . . .	31
3.2.1 Biografisches . . . . .	31
3.2.2 Leonhard Euler und das » <i>Tentamen</i> « . . . . .	32
3.2.3 Die »Vokalorgel« . . . . .	43
3.2.4 Zusammenfassung . . . . .	51
3.3 Abbé Mical . . . . .	52
3.4 Wolfgang von Kempelen . . . . .	58
3.4.1 Historische Aspekte der Sprachmaschine . . . . .	58
3.4.2 Repliken von ca. 1800 bis heute . . . . .	68
3.5 Zusammenfassung . . . . .	72

<b>4</b>	<b>19. Jh.: Epigonen &amp; Erben</b>	<b>75</b>
4.1	Letzte Pseudo-Sprachmaschinen . . . . .	75
4.1.1	L. Müller . . . . .	75
4.1.2	Heinrich Markus Brunner . . . . .	78
4.2	In Kempelens Schatten . . . . .	79
4.2.1	Carl Salomon Warmholz . . . . .	79
4.2.2	Leonhard Posch . . . . .	81
4.3	Die geistigen Erben . . . . .	83
4.3.1	Robert Willis . . . . .	83
4.3.2	Sir Charles Wheatstone . . . . .	86
4.3.3	Joseph Faber . . . . .	89
4.3.4	Melville und Alexander Graham Bell . . . . .	96
4.4	Zwischen Wissenschaft und Sensation . . . . .	97
4.4.1	Johann Nepomuk Mälzel . . . . .	98
4.4.2	Die thüringischen »Stimmenmacher« . . . . .	100
4.5	Zusammenfassung . . . . .	101
<b>5</b>	<b>20. und 21. Jh.: Letzte Innovationen</b>	<b>103</b>
5.1	Sprachsynthesen der »Bell Laboratories« . . . . .	103
5.2	Vokalsynthesen mit Labialpfeifen . . . . .	106
5.2.1	Dayton Clarence Miller . . . . .	106
5.2.2	G. Oscar Russel & Arthur Taber Jones . . . . .	110
5.3	Vokalsynthese mit Vokaltraktmodellen . . . . .	111
5.4	»Sprechende Köpfe« im 21. Jahrhundert . . . . .	111
5.4.1	Die »Waseda Talkers« . . . . .	113
5.4.2	Der »Furhat Social Companion Talking Head« . . . . .	116
5.5	Exkurs: Ein »sprechendes Klavier« . . . . .	116
5.6	Zusammenfassung . . . . .	118
<b>II</b>	<b>Historische Sprachsynthesen im Experiment</b>	<b>123</b>
<b>6</b>	<b>Vox humana – Eine Vokalsynthese?</b>	<b>125</b>
6.1	Frühere akustische Analysen . . . . .	125
6.2	Eigene akustische Analysen . . . . .	128
6.2.1	Datengrundlage . . . . .	128
6.2.2	Analyse . . . . .	134



6.3	Perzeptionsexperimente . . . . .	143
6.3.1	Fragestellungen und Methoden . . . . .	143
6.3.1.1	Experiment Nr. 1 . . . . .	143
6.3.1.2	Experiment Nr. 2 . . . . .	144
6.3.1.3	Experiment Nr. 3 . . . . .	145
6.3.2	Ergebnisse . . . . .	147
6.4	Diskussion . . . . .	155
<b>7</b>	<b>Kratzensteins »Vokalorgel«</b>	<b>163</b>
7.1	Die durchschlagende Lingualpfeife . . . . .	163
7.2	Zur möglichen Rekonstruktion der »Vokalorgel« . . . . .	164
<b>8</b>	<b>Kempelens Sprachmaschine</b>	<b>169</b>
8.1	Die Saarbrücker Repliken . . . . .	169
8.2	Alternative Anregungskonzepte . . . . .	170
8.2.1	Kempelens aufschlagende Lingualpfeife . . . . .	170
8.2.2	Varianten der aufschlagenden Lingualpfeife . . . . .	171
8.2.3	Durchschlagende Zungenpfeifen . . . . .	173
8.2.4	Zachariaszungen . . . . .	179
8.2.5	Doppelrohrblätter . . . . .	180
8.2.6	Eine membranöse Pfeife . . . . .	183
8.3	Perzeptionsexp. mit Sprachmaschinen-Repliken . . . . .	184
8.3.1	Ausgangslage . . . . .	184
8.3.2	Datengrundlage und Durchführung . . . . .	185
8.3.3	Ergebnisse . . . . .	186
8.3.4	Diskussion . . . . .	187
<b>III</b>	<b>Resumée</b>	<b>191</b>

Danksagung	197
Tabellenverzeichnis	XI
Abbildungsverzeichnis	XIII
Literaturverzeichnis	XVII
Index	LII
Appendix	LVII
Auszüge aus Kratzensteins "Tentamen"	LVII
Glossar	LXII
Audio- und Video-Material	LXXI

# Einleitung

»Kann was natürlicher, als Vox humana, klingen?« fragte 1733 WILHELM FRIEDEMANN BACH (1710–1784) in Bezug auf die Orgel der Dresdner Frauenkirche. Eine Frage, die im Zusammenhang mit Sprachsynthese aus heutiger Sicht zunächst nicht als thematisch zugehörig erscheint. Heute ist die – elektronische – Generierung synthetischer Sprache und deren Nutzung alltäglich. Wenngleich die Synthesequalität oft immer noch mehr oder weniger stark zu wünschen übrig lässt und praktisch keine Synthese für sich in Anspruch nehmen kann, vollkommen authentisch und »natürlich« zu wirken, so scheint mit dem aktuellen Status quo doch erstmals das Ziel erreicht zu sein, von dem eine Vielzahl von Menschen – auch aus ganz verschiedenen Gründen – seit Jahrhunderten träumte: Spracherzeugung unabhängig von einem Sprecher ist alltäglich geworden und dient als mannigfaltiges Hilfsmittel, sei es als Ansagesystem auf Bahnhöfen, sei es als lebensunterstützende »Prothese« für Blinde oder Menschen mit die Sprache behindernden Einschränkungen wie – um nur ein prominentes Beispiel zu nennen – STEPHEN HAWKING (\*1942).<sup>1</sup>

Wohl nur wenige aber wissen darum, wie lange einerseits schon an dieser Problemstellung geforscht wird und andererseits, wie häufig sich Forscher immer wieder kurz vor dem Durchbruch zu einer praxistauglichen Synthese wähnten, nur um dann doch an als grundlegend erkannten Schwierigkeiten (die meist dieselben blieben) zu scheitern.

Dies liegt wohl auch daran, dass wir Sprachsynthese heute nahezu ausschließlich als elektronisch via Software erzeugt erleben und dass aus dieser Perspektive die Versuche, mit rein mechanischen Lösungen zum Ziel zu kommen, wahrscheinlich abenteuerlich wirken müssen. Das mechanische Konzept jedoch ist nicht nur mehrere Jahrhunderte alt – legendenhafte

---

<sup>1</sup> STEPHEN HAWKING nutzt im Übrigen nach Auskunft seiner Tochter LUCY HAWKING bis zum heutigen Tage eine Klatt-Synthese aus dem Jahre 1984, die lediglich hinsichtlich ihrer Steuerung modernisiert wurde. HAWKING betrachtet diese Synthese heute als seine persönliche Stimme.

Berichte reichen bis ins frühe Mittelalter – es ist auch heute noch lange nicht ad acta gelegt. Immer noch gibt es vereinzelte Versuche, Sprache auf mechanischem Wege zu synthetisieren, wenn auch freilich nun mit massiver Unterstützung durch Computer.

Der lange und kurvenreiche Weg der mechanischen Sprachsynthese mit seinen zahlreichen in Sackgassen endenden Verzweigungen, von den legendären Anfängen bis zu den robotischen Konzepten der Gegenwart, soll das zentrale Thema dieser Arbeit sein. Der Schwerpunkt wird hierbei auf dem Zeitraum des 18. und 19. Jahrhunderts liegen, in dem eine besonders intensive Auseinandersetzung mit dem Thema stattfand. Genannt seien hier besonders die »Stars« der mechanischen Sprachsynthese, WOLFGANG VON KEMPELEN (1734-1804) und CHRISTIAN GOTTLIEB KRATZENSTEIN (1723-1795).

Gerade diese beiden stehen aber auch für die aus heutiger Sicht absonderlichen Querverbindungen und Wechselbeziehungen, die die Sprachsynthese mit thematisch scheinbar völlig anderen Themengebieten wie dem Instrumentenbau einging. Insbesondere dem Orgelbau wird hier ein intensives und ausführliches Augenmerk zu widmen sein, da er die Grundlage für nahezu alle in dieser Arbeit vorgestellten Synthesekonzepte des 18. Jahrhunderts und darüber hinaus bildet.

Zwar scheint bis zum heutigen Tage kein Prototyp einer mechanischen Sprachsynthese auch langfristig tatsächlich erfolgreich gewesen zu sein. Trotzdem ist es faszinierend, sich mit der Genese zu beschäftigen, die man jedenfalls für den Zeitraum bis weit in das 19. Jahrhundert hinein als »Variationen über ein gegebenes Thema« bezeichnen könnte, nämlich die einmal entstandene und sich dann hartnäckig über sehr lange Zeit haltende Vorstellung, in der Orgel das unmittelbare Vorbild für eine effektive Sprachsynthese zu besitzen. Ein resumierendes Zitat DAVID BREWSTERS (1781–1868) in seinen »*Briefen über die natürliche Magie*« zeigt, wie nahe man sich spätestens zu diesem Zeitpunkt – der Mitte des 19. Jahrhunderts – einer wirklichen Lösung glaubte:

»[...] Es leidet wohl keinen Zweifel, daß, ehe noch ein Jahrhundert verflossen seyn wird, die Wissenschaft sprechende und singende Maschinen ihren Bereicherungen beizählen werde.«  
[Brewster 1833, 250]

BREWSTERS Prognose sollte sich zwar als zutreffend erweisen, doch in einer vollständig anderen Art (nämlich als rein reproduzierende Maschinen), als er vermutet haben dürfte. Wann erstmals die Idee aufkam, *Zungenpfeifen* – genauer gesagt solche der *Regalbauweise*, vgl. Kapitel 2.2 – als Anregungskonzept zu verwenden, ist unbekannt. Aus heutiger Perspektive wirkt die

Idee zumindest kurios, ein System zur Sprachsynthese von einem Musikinstrument ableiten zu wollen. Doch ist es offenkundig, dass schon früh eine Wechselbeziehung zwischen der *Vox humana* der Orgel und den Konzepten zur mechanischen Sprachsynthese bestand (vgl. Kapitel 2 und 3). Auch das für den Titel verwendete Zitat WILHELM FRIEDEMANN BACHS zeugt – wenn auch aus Sicht eines Musikers – von dieser Vorstellung.

Bislang liegt keine Arbeit vor, die sowohl einen ausführlichen geschichtlichen Überblick über die Entwicklung der mechanischen Sprachsynthese seit dem 18. Jahrhundert bietet als auch die technologischen Konzepte dieser Epoche in Zusammenhang setzt mit ihren unstreitigen Vorbildern im Orgelbau. Dies ist durchaus verständlich, bedarf es dazu doch umfassender Kenntnisse in den beiden recht unterschiedlichen Bereichen Phonetik und Orgelbau und hier insbesondere in den jeweiligen historischen Zusammenhängen. Einen ausführlichen Abriss der Sprachsynthese-Geschichte – unter Umgehung der orgelbaulichen Thematik – bietet [Köster 1973], der für die Konzeption dieser Arbeit und die damit verbundenen Recherchen eine wertvolle Hilfestellung leisten konnte. Insbesondere das wenig verbreitete Wissen um konstruktive und historische Aspekte des Orgelbaues führte auch dazu, dass bisweilen die engen Verbindungen der frühen Sprachsyntheseforschung mit dem Orgelbau nicht adäquat bewertet oder gar nicht erst erkannt wurden. Frühere Arbeiten des Autors konnten dieses vielfältige Themengebiet zwar bereits grob umreißen [vgl. Brackhane & Trouvain 2013b], eine ausführliche Behandlung steht indessen noch aus.

Die vorliegende Arbeit will dieses Forschungs-Desiderat beseitigen und die maßgebliche Rolle, die der Orgelbau im Allgemeinen und eine bestimmte Pfeifenbauform im Besonderen in der Sprachsyntheseforschung seit dem 18. Jahrhundert spielten, angemessen darstellen. Hierzu gehört auch eine kritische experimentalphonetische Auseinandersetzung mit eben jener Pfeifenbauform, der *Vox humana*.

Diese Arbeit ist in zwei Hauptabschnitte gegliedert: Zunächst soll ein historischer Abriss der mechanischen Sprachsynthese entworfen werden. Er soll deutlich machen, wie beherrschend die frühzeitig als Irrweg erkannte Theorie von dem Orgelregister *Vox humana* in der Praxis als Anregungskonzept trotzdem bis ins 20. Jahrhundert hinein blieb. Zum anderen soll dieser Abschnitt auch den Einfluss der Synthesekonzepte WOLFGANG VON KEMPELENS und CHRISTIAN GOTTLIEB KRATZENSTEINS aufzeigen, von denen sich die Forschergemeinde erst im 20. Jahrhundert weitenteils zu lösen vermochte. Im zweiten Abschnitt werden zunächst phonetische Analysen und Perzeptionstests zu verschiedenen Aspekten der mechanischen Sprachsynthese und der Vokalähnlichkeit der *Vox humana* vorgestellt. Hierbei soll untersucht werden, welche akustischen Parameter für

die weitgehende Akzeptanz der Klänge einer *Vox humana* und der Synthesekonzepte KEMPELENS und KRATZENSTEINS verantwortlich sind. Die folgenden Kapitel widmen sich der möglichen Rekonstruktion bzw. Modifikationen der Synthesekonzepte KRATZENSTEINS und KEMPELENS.

Die umfangreicheren Kapitel 2–5 des historischen Teils I enden jeweils mit einer knappen Zusammenfassung, zusätzlich auch der komplexe Abschnitt 3.2 über CHRISTIAN GOTTLIEB KRATZENSTEIN und seiner Beziehung zu LEONHARD EULER. Der zusammenfassende Abschnitt 5.6 leitet zugleich auch als kritische Würdigung von der in Teil I dargestellten Entwicklung über zu den hieraus resultierenden experimentalphonetischen Kapiteln des Teil II.

Bei den Recherchen zum historischen Teil dieser Arbeit wurde immer deutlicher, dass die Sprachsynthesekonzepte KRATZENSTEINS und KEMPELENS eine Besonderheit darstellen: Zu ihnen existieren – wenn auch für KRATZENSTEINS »Vokalorgel« nur mit deutlichen Einschränkungen – Beschreibungen aus den Händen ihrer Schöpfer, anhand derer man sich eine recht genaue Vorstellung der Konstruktionen machen kann. Für alle anderen hier behandelten *Sprachmaschinen* des 18. und 19., ja sogar viele des 20. Jahrhunderts, gilt dies leider nicht, so dass sich ihre Beschreibung gezwungenermaßen auf Berichte Dritter stützen muss, die naturgemäß in ihrer Präzision und Detailliertheit weit hinter Darstellungen aus erster Hand zurückbleiben müssen.

Insbesondere bei den Nachforschungen über den Verbleib der zahlreichen *Sprachmaschinen* bzw. ihrer Repliken des 19. Jahrhunderts wurden die immensen, besonders in Deutschland zum großen Teil kriegsbedingten, Verluste an Exponaten und Akten der betreffenden Museen und Archive schmerzhaft deutlich. Zahlreiche zunächst vielversprechende Spuren verloren sich wegen fehlender Dokumentationen im Nichts. Der letztendliche Nachweis über den tatsächlichen Verlust oder die Existenz der Maschinen wäre in allen Fällen nur unter Aufbietung eines immensen Zeitaufwandes durch intensive Recherchen in den meistens nicht digitalisierten, teilweise sogar ungeordneten Altbeständen der zuständigen Archive möglich gewesen und musste daher leider vorerst unterbleiben.

## Anmerkungen zum Quellengebrauch

Anders als sonst üblich wurde darauf verzichtet, die im historischen Teil dieser Arbeit verstärkt auftretenden teils umfangreichen Zitate paraphrasierend zusammenzufassen. Dies hätte zwar einerseits den Vorteil der deutlichen Platzersparnis, ginge aber andererseits insbesondere bei historischen (Primär-)Quellen zulasten der in einem Text immer auch enthaltenen Me-

tainformationen. Die zum näheren Verständnis der Konstruktion der »Vokalorgel« CHRISTIAN GOTTLIEB KRATZENSTEINS (vgl. Kap. 3.2 und 7) notwendigen, allerdings recht umfangreichen Zitate aus seiner Schrift »*Tentamen resolvendi problema...*« finden sich in einem separaten Appendix (S. LVII), auf den an den notwendigen Stellen jeweils verwiesen wird.

Zitate in französischer und lateinischer Sprache werden – nicht zuletzt aufgrund der häufig auftretenden organologischen Fachtermini – zur leichteren Verständlichkeit jeweils sowohl im originalen Wortlaut als auch in einer deutschen Übersetzung angeführt, die jeweils – sofern nicht ausdrücklich anders gekennzeichnet – von mir selbst stammt.<sup>2</sup> Diese Übersetzung ist jeweils typografisch so abgesetzt, dass sie als sekundäre Zutat deutlich wird. Für englische Zitate wurde auf Übersetzungen verzichtet, da bei ihnen von einer allgemeinen Verständlichkeit ausgegangen werden kann.

Die Verwendung von auch heute noch gebräuchlichen Fachtermini in historischen Quellen stellt eine besondere Problematik dar: Oft besaßen diese seinerzeit nicht exakt dieselbe Bedeutung wie heute oder sie wurden in ihrer Verwendung eher lax gehandhabt, was schlimmstenfalls zu grundlegenden Missverständnissen führen kann. Dies gilt insbesondere – aber nicht ausschließlich – für die orgelbaulichen Termini *technici* in den Texten CHRISTIAN GOTTLIEB KRATZENSTEINS. Hier bedurfte es stets einer kritischen Überprüfung.

Die im Literaturverzeichnis mit einem Asterisken gekennzeichneten historischen Quellen konnten via Internet in digitalen Bibliotheken wie *Google Books* ausgewertet werden, so dass keine aufwändige Fernleihe notwendig war.

Für die insbesondere im ersten Teil dieser Arbeit verwendeten zahlreichen historischen Quellen wurden die jeweiligen Einträge im Literaturverzeichnis mit erläuternden Kommentaren versehen. Soweit möglich, wurden Informationen zum Autor und der Kontext der jeweiligen Veröffentlichung bzw. des verwendeten Zitates angegeben.

Einen besonderen Stellenwert innerhalb der verwendeten Literatur nimmt die Gattung des *Offenen Briefs* ein, von der im 18. und 19. Jahrhundert reger Gebrauch gemacht wurde, so dass gleich acht der hier verwendeten historischen Quellen diesem Genre angehören. Zur Geschichte des Offenen Briefs als Medium von Gelehrten und Kritikern vgl. [Essig 2000]. Hervorzuheben ist, dass Absender und/oder Adressat bei dieser Gattung durchaus auch fingiert sein konnten (vgl. [Essig 2000, 117]).

---

<sup>2</sup>Für die kritische Durchsicht und Verfeinerung meiner Übersetzungen habe ich Frau JACQUELINE KUBCZAK (französische Texte) und Frau Dr. ELISABETH LINK (lateinische Texte) herzlich zu danken.

Den wissenschaftlichen Gepflogenheiten folgend wurde auf die Verwendung von Sekundärzitataten soweit als irgend möglich verzichtet und stattdessen jeweils die genannte Primärquelle aufgesucht und referenziert. Nur in Einzelfällen, in denen sich die (historische) Originalquelle nicht mit einem noch vertretbaren Rechercheaufwand ermitteln ließ, die fragliche Information jedoch zugleich als wichtig und glaubwürdig eingeschätzt wurde, gelangten Sekundärzitate zur Anwendung.

## Anmerkungen zur Typografie und den Appendices

Die innerhalb von Zitaten stehenden Kommentare/Erläuterungen in eckigen Klammern sind stets Ergänzungen des Autors und nicht Teil des Zitats. Die Angabe »[sic!]*«* verweist hierbei auf missverständliche orthografische, die Angabe »[!]*«* auf bemerkenswerte inhaltliche Aspekte. Umfangreichere Kommentare wurden als Fußnoten gesetzt. Eckige Klammern finden auch für Quellenangaben Verwendung. Im laufenden Text stehen Ergänzungen wie z. B. Lebensdaten grundsätzlich in runden Klammern.

Für die Bezeichnung von Vokalqualitäten wurden parallel drei Systeme verwendet: Sofern nicht von phonetisch definierten Qualitäten, sondern auf Grundlage der Orthografie gebildeten Vokalkategorien die Rede ist, werden diese Vokale auch rein orthografisch als Majuskeln wiedergegeben, also A, E, I, O, U. Eine phonetisch weite, kategorielle Qualifizierung wurde durch // ausgedrückt: /a/, /e/ etc. Vokalsymbole in eckigen Klammern bezeichnen schließlich phonetisch eng definierte Vokalqualitäten: [a], [i] etc.

Im Text genannte Titel von Büchern oder Artikeln werden stets im Format »*Titel*« angegeben. Auftretende termini technici wurden *kursiv* gesetzt. Sofern es sich hierbei um spezifische Begriffe aus dem Orgelbau handelt, wie sie insbesondere in Kapitel 2 verstärkt auftreten, werden sie im Glossar ab Seite LXII erläutert.

Ergänzend wurde eine Sammlung von Audio-Beispielen zusammengestellt, die zum einen Hörbeispiele zur musikalischen Verwendung der *Vox humana* im 18. Jahrhundert enthält und zum anderen sämtliche Aufnahmen bzw. Stimuli, die den Untersuchungen der Kapitel 5 + 6 zu Grunde liegen. Nähere Erläuterungen hierzu finden sich im Appendix ab Seite LXXI.

Ebenfalls vorhanden sind ein Stichwort- sowie separate Tabellen- und Abbildungsverzeichnisse.



## Anmerkungen zur Terminologie

Auch wenn es faktisch nur für einen Zeitraum weniger Jahrzehnte von tatsächlicher Bedeutung ist, so gilt es doch, die unterschiedlichen, meist quasi-synonym verwendeten Termini zur mechanischen Sprachsynthese zu differenzieren. Bis in das 19. Jahrhundert hinein werden die Bezeichnungen *Sprechende Maschine*, *Sprechmaschine* und *Sprachmaschine* synonym gebraucht. Erst mit dem Aufkommen der Phonographen ändert sich die Sprachregelung. Da man Phonographen als die moderne Weiterentwicklung der historischen mechanischen Sprachsynthesen begriff, wurden auch sie oftmals als *Sprechmaschinen* bezeichnet, nie allerdings als *Sprachmaschinen* (vgl. bspw. [Reko 1911] und [Niemann 1920]). Dies ist auch insofern konsequent, als Phonographen zwar zu sprechen scheinen, nie aber selbstständig Sprache generieren.

Auf diesen Umstand aufmerksam geworden, erkennt man bald, dass für den hier behandelten Themenkomplex die Bezeichnung *Sprachmaschine* in der Tat zutreffender ist, da er sich definatorisch weiter fassen lässt: Bei den »klassischen« Konstruktionen der mechanischen Sprachsynthese handelt es sich – zumindest theoretisch – um Vorrichtungen, die in der Lage sind, Sprache frei, d.h. ohne Repertoire-Einschränkung durch ein wie auch immer geartetes Speichermedium zu produzieren, wohin gegen sich *Sprechmaschinen* als Vorrichtungen zur bloßen Reproduktion von Sprache (d.i. Phonographen) definieren ließen. Eine Ausnahme dürfte allerdings die Konstruktion MICALS sein (siehe 3.3), die offenbar – zumindest auch – auf Grundlage eines *Walzenmechanismus* ähnlich dem einer Spieluhr funktionierte, also nach der hier getroffenen Definition sowohl *Sprach-* als auch *Sprechmaschine* war. Ein weiterer Unterschied besteht zudem darin, dass *Sprechmaschinen* im Prinzip ohne ständige und direkte Kontrolle durch einen Menschen funktionieren, *Sprachmaschinen* hingegen prinzipiell in Echtzeit von einem Menschen gesteuert werden müssen.



# Teil I

## Mechanische Sprachsynthese – ein historischer Überblick



# Kapitel 1

## Frühe Berichte zu Sprachsynthesen

Bereits aus der Antike sind Berichte über sog. »Sprachmaschinen« überliefert. Allerdings wird man wohl allen derartigen Erfindungen bis mindestens um das Jahr 1700, in mehreren Fällen aber auch bis weit in das 18. Jahrhundert hinein, das Prädikat der »Pseudo-Sprachmaschine« verleihen müssen. Zwar erschöpft sich die Überlieferung nur allzu oft in legendenhaften Erzählungen ohne präzise und verlässliche Angaben, so dass es nahezu unmöglich ist, den jeweiligen individuellen Tatsachengehalt einer Überprüfung zu unterziehen. Doch lassen sowohl diese Beschreibungen als auch rückblickende Berichte aus dem 17. und 18. Jahrhundert erkennen, dass man die früheren Konstrukte allgemein für bessere Taschenspielertricks hielt. Eine Übersicht über die seinerzeit legendenhaft bekannten »Sprachmaschinen« bietet BUSCHS »*Handbuch der Erfindungen*« [Busch 1821, 482 f.].

Nach Lage der Quellen handelte es sich bei allen diesen »Sprachmaschinen« um zumindest oberflächlich, also äußerlich mehr oder weniger anatomisch korrekt modellierte humanoide Figuren oder Köpfe, deren geöffnetem Mund in irgendeiner Form durch einen Menschen erzeugter Schall zugeführt wurde [Niemann 1920, 2 f.]. So darf man auch bei den für ALBERTUS MAGNUS (~1200–1280) oder ROGER BACON (1214–1292?) überlieferten »Sprechenden Köpfen« annehmen, dass es sich um von der Idee des Golems inspirierte sagenhafte Objekte handelte, deren historischer Kern höchstens in einem für Zeitgenossen »mysteriösen« Interesse an der Nachvollziehung physiologischer Prozesse zu suchen wäre (vgl. [Wheat-

stone 1879, 348 ff.] und [Niemann 1920, 4 f.] sowie [Jackson 2005, 13]). Ein Bericht über Gestalt und Funktion einer derartigen »sprechenden Maschine« findet sich im »*Hannoverschen Magazin*« von 28. Dezember 1767 [Anonymus 1767, 1649 ff.].

In allen diesen Fällen handelte es sich also um Versuche, die menschliche Anatomie nachzubilden und erst in einem weiteren Schritt durch die Ingangsetzung dieser anatomischen Konstrukte das akustische Resultat zu erhalten, also Sprache. Dies sollte in der gesamten Geschichte der mechanischen Sprachsynthese die bestimmende Konstante bleiben. Versuche, Sprachsynthese unter Umgehung der physiologisch-artikulatorischen Nachahmung zustande zu bringen, blieben in der Minderheit.

Noch Jahrzehnte später – de facto bis in KEMPELENS (1734–1804) Zeiten – hatte die Idee von den »Sprechenden Köpfen« nichts an Faszination eingebüßt. Auch für den Jesuiten und Universalgelehrten ATHANASIUS KIRCHER (1602–1680) sind Berichte über solche Objekte überliefert. REKO zufolge hatte Kircher in seinem römischen »Museo Kircheriano« mehrere

»sprechende Büsten und Köpfe aufgestellt. Dieselben hatten geheime Zuführungsröhren (sogenannte Sprachrohre) zu ihrem Munde. Fragte man einen der Köpfe etwas, so konnte ein im Keller versteckter Mann durch Zuführungsröhren die Frage hören und eine geschickte und passende Antwort nach seinem Belieben darauf erteilen, die dann so klang, als würde der betreffende Kopf selbst sprechen. [...] Athanasius Kirchner [sic!] bezeichnete übrigens selbst diese seine sprechenden Köpfe als bloße Vorläufer einer wirklich sprechenden Maschine, an der er seit langem gearbeitet haben will. Die Grundideen, auf denen diese universale Sprechmaschine hätte beruhen sollen, hat er in seiner *Musurgia universalis*, die 1650 erschien, niedergelegt [vgl. Kircher 1650, 305 ff. und Kircher 1684, 81 f.].« [Reko 1911, 415]

Eine ähnliche Konstruktion dürfte dem »Sprechenden Kopf« des Dresdner Gelehrten und Konrektors der Kreuzschule JOHANN VALENTIN MERBITZ (1650–1704) zugrunde gelegen haben, über den unlängst RÜDIGER HOFFMANN erstmals eingehend berichtet hat [Hoffmann 2009]. Der am einfachsten zugängliche Bericht entstammt der »*Allgemeinen Deutschen Biografie*«, ist demnach mit einem Abstand von annähernd 200 Jahren entstanden.<sup>3</sup>

<sup>3</sup>Dieser Artikel fußt allerdings seinerseits auf der Leipziger Schrift *De loquela imaginum* des CHRISTIAN FLEMING (1705). Frdl. Mitteilung von Prof. Dr. RÜDIGER HOFFMANN.

»Das Merkwürdigste jedoch, was über ihn berichtet wird, ist, daß er in fünfjähriger Arbeit einen kunstvollen Kopf gebildet haben soll, der auf jede Art von Fragen, die man ihm ins Ohr sagte, mit deutlicher Stimme und in allen Sprachen, auch lateinisch, französisch, hebräisch und griechisch, zu antworten vermochte, Zukünftiges vorhersagte und Geheimes offenbarte. Nur durch den Tod soll M. verhindert worden sein, ein noch kunstreicheres Werk zu vollenden, mit dem er bereits acht Jahre beschäftigt war: er beabsichtigte zwei Bildsäulen herzustellen, die mit einander über jeden beliebigen Gegenstand ein Zwiegespräch in Frage und Antwort zu führen und dabei jeden gewünschten guten oder übeln Geruch, Zimmet- oder Rosen- oder Myrrhengeruch u. s. w., von sich zu geben verstanden.« [Schnorr von Carolsfeld 1885]

Zuverlässige zeitgenössische Berichte waren nicht aufzufinden. Es steht indessen zu vermuten, dass es sich hier um eine der Kircherschen sehr ähnliche Vorrichtung gehandelt haben dürfte.

Eine andere Form des »Sprechenden Kopfes« entwickelte der Naturforscher und Leibmedikus des schwedischen Königs Karl XI., URBAN HJÄRNE (1641–1724). Während eines Frankreichaufenthaltes 1670 befasste er sich mit der Problemstellung der Sprachsynthese und unternahm hierzu Versuche mit den Köpfen und speziell Laryngen von menschlichen Leichen.<sup>4</sup>

CHARLES WHEATSTONE (1802–1875) beschreibt mehrere »Sprechende Köpfe«, die beispielsweise an den Höfen von London und Versailles für Aufsehen sorgten, hinter denen sich aber bei genauerer Betrachtung ein menschlicher Sprecher verbarg. [Wheatstone 1879, 350 f.] NIEMANN zufolge scheint es wenige Jahrzehnte nach KIRCHER einen ersten Schritt in die später intensiv verfolgte Richtung gegeben zu haben:

»Die Vervollkommnung der Orgel, besonders die Vermehrung der Stimmen [Register], wies nun den Weg, den man bei der Konstruktion von Sprechapparaten einzuschlagen hatte, zeigte doch die vox humana, daß es möglich war, dem Klang der menschlichen Stimme nahe zu kommen. Freilich, von dem einfachen Laut bis zur artikulierten Sprache war noch ein weiter Weg. Wichtig aber war es, daß man von nun an auch in wissenschaftlichen Kreisen dem Problem Beachtung schenkte. Samuel Reyher [1635–1714, Deutscher Mathematiker und Astroном] war wohl der erste, der sich wenigstens theoretisch mit

---

<sup>4</sup>Freundliche Mitteilung von Mats Ljungquist.

der Sprechmaschine befaßte<sup>5</sup>. Er meinte die Vokale und Verbindungen von Vokalen und Konsonanten müßten sich durch einzelne Orgelstimmen wiedergeben lassen, und durch gleichzeitiges Anschlagen einiger Tasten müßte man eine Mischung und Aneinanderreihung von Tönen erhalten [Reyher 1693, 12].<sup>6</sup>

Einem unbekannten Erfinder scheint es nun in der Tat gelungen zu sein, wenigstens eine beschränkte Anzahl von Lautverbindungen in der oben angedeuteten Weise wiederzugeben. Georg Pasch [(1661–1707), deutscher Theologe] [...] berichtete nämlich, er habe in Paris (Ende des 17. Jahrh.) eine hölzerne Statue gesehen, die mit Hülfe einer Luftorgel deutliche Sätze, aber nur in beschränkter Anzahl sprach.« [Niemann 1920, 18 f.]

Hier scheint der gedankliche Weg zu der ein knappes Jahrhundert später durch LEONHARD EULER skizzierten Verfahrensweise nur mehr kurz zu sein (vgl. Kapitel 3.2.2). Indessen stellt sich angesichts der erwähnten »hölzernen Statue« sogleich die Frage, inwieweit bei dieser Apparatur alles mit rechten Dingen zugeht.

Nun tut sich ein beinahe hundert Jahre dauerndes Überlieferungsloch auf. Erst um 1770 erhalten wir wieder ausführliche und ernstzunehmende Nachrichten über Sprachsynthesekonzepte, bei denen nicht von vornherein angenommen werden muss, dass es sich um ein Konstrukt kircherscher Art handelte. Um die Mitte des 18. Jahrhunderts unternehmen beinahe zeitgleich und vollkommen unabhängig voneinander zumindest vier Forscher den Versuch, eine zumindest mit Einschränkungen praxistaugliche mechanische artikulatorische Sprachsynthese zu erschaffen. Gerade im Hinblick zu dem vorangehenden Vakuum auf diesem Gebiet erscheint es so, als habe die Entwicklung geradezu in der Luft gelegen.

In seiner Abhandlung »*La rhétorique ou l'art de parler*« entwickelte der französische Philosoph und Mathematiker BERNARD LAMY (1640–1715) ein Gedankenexperiment, das die praktische Unmöglichkeit einer Sprachmaschine beweisen sollte [Lamy 1699, 153 ff.]. Diese Unmöglichkeit dient ihm zugleich als Beweis für den göttlichen Ursprung der Sprache, denn nur Gott selbst könne etwas so Komplexes und Unwahrscheinliches Wirklichkeit werden lassen.

Als Basis für seine Überlegungen bedient sich auch LAMY einmal mehr der Hypothese, dass die Orgel zumindest theoretisch alle Voraussetzun-

<sup>5</sup>Vgl. [Reyher 1693, 12]. Dass Reyher indessen bei weiten nicht der erste war, der sich mit dem Thema befasst hat, wurde oben dargelegt.

<sup>6</sup>Eine ähnliche Idee wurde vor wenigen Jahren vom Komponisten PETER ABLINGER in Form eines »Sprechenden Klaviers« verwirklicht, siehe Kap. 5.5.



gen für eine Sprachsynthese biete: je nach Konstruktionsart der einzelnen Pfeifen sei es möglich, die verschiedensten Laute zu erzeugen. Wenn es gelänge, für jeden Laut der menschlichen Sprache eine passende Pfeife zu konstruieren, hätte man eine sprechende Orgel oder Sprachmaschine. Für das Französische mit seinen fünf Vokalen und siebzehn Konsonanten seien also 22 Pfeifen notwendig.

Indessen bestünde Sprache laut LAMY aus mehr als der bloßen Aneinanderreihung von Einzellauten. Ein »a« und ein »b« aneinandergereiht, erbrächten daher noch lange kein »ab«. Also sei es notwendig, für jede nur denkbare Lautkombination innerhalb einer Silbe eine eigene Maschine zu bauen, was zu einer ungeheuer großen Anzahl von Einzelgeneratoren führe.

Das, was der Mensch also lediglich mithilfe seines Mundes, seiner Zunge, seiner Lippen und einer Reihe von Muskeln vollbringe, könne demnach auf mechanischem Wege niemals nachgeahmt werden. Nur jene, die nichts darüber wüssten, wie die menschliche Sprache funktioniere, könnten glauben, es sei jemals möglich, den menschlichen Sprechapparat nachzubauen. Zwei Generationen nach dieser Niederschrift wurde gleich mehrfach der Beweis dafür angetreten, dass LAMY zumindest in seiner Rigidität irrte.



# Kapitel 2

## Das Orgelregister »Vox humana«

Zu den im Folgenden auftretenden *kursiv* gesetzten orgelbaulichen Fachtermini vgl. das Glossar ab S. LXII.

### 2.1 Etymologie und Geschichte

Scheinen auch Orgelbau und Sprachsynthese zunächst nichts miteinander gemeinsam zu haben, so geben die im vorangegangenen Kapitel skizzierten Theorien einen Eindruck davon, dass man im 17. und 18. Jahrhundert sehr wohl versuchte, einen Zusammenhang zwischen diesen beiden Gebieten herzustellen. Es mag auf den ersten Blick durchaus verwundern, dass eine Pfeifenbauform aus dem Orgelbau einmal zum Gegenstand, wenn nicht gar zur *idée fixe* der Forschung zur mechanischen Sprachsynthese besonders im 18. Jahrhundert werden sollte. Tatsächlich scheint es sich hierbei um die geradezu tragikomische Verkettung mehrerer Missverständnisse und Fehleinschätzungen zu handeln, indem zum einen die Orgel als komplexer, aber doch verhältnismäßig einfach zu steuernder Mechanismus als Prototyp der Steuerung einer mechanischen Sprachsynthese angesehen (bspw. [Dodart 1700, 239 ff.]) und zum anderen die Bezeichnung *Vox humana* (im weiteren Verlauf dieser Arbeit stets als VH bezeichnet) häufig als dezidierte Klangbeschreibung verstanden wurde und nicht als eine eher abstrakte Umschreibung.

Das Orgelregister VH gehört zu den sog. *kurzbecherigen Lingualregistern* oder *Regalen*. Gemeinsames Merkmal aller *Regale* ist, dass die Länge

des *Resonators* einer Pfeife in keinem direkten Zusammenhang mit ihrer jeweiligen Tonhöhe steht, sondern recht kurz ist und über Intervalle hinweg relativ konstant bleibt. Für Pfeifen der *Regalbauart* ist der *Resonator* bei VH-Pfeifen jedoch mit einer Länge von maximal 1' auf C verhältnismäßig lang. Hinzu kommt, dass die *Resonatoren* der VH stets *teilgedeckt* sind, was eine gewisse Dämpfung des Klanges bewirkt. Die VH entstand in der Mitte des 16. Jahrhunderts in Frankreich mit einer genau definierten Bauform, wie sie später in Deutschland beispielsweise durch die Gebrüder STUMM tradiert wurde (s. u.). Die VH wurde hier selten solistisch, sondern zumeist in der Kombination mit einem grundtönigen *gedeckten 8'-Register* (*Bourdon* bzw. *Gedeckt*) und ggf. dem *Tremulanten* gespielt.<sup>7</sup> Nachdem sie einige Jahrzehnte zuvor schon in den Niederlanden großen Anklang gefunden hatte, wurde die VH ab ca. 1650 auch in Deutschland populär, wo sich alsbald eine geradezu unüberschaubare Vielfalt an Bauformen ausbildete [Busch 2007, 817].

Auch deutsche Musiktheoretiker waren sich spätestens seit dem Auftreten der VH in Deutschland im 17. Jahrhundert der eigenartigen Faszination dieses Orgelregisters bewusst und der Problematik, die sich hieraus ergab. Dies zeigt sich besonders häufig in Beschreibungen dieser Pfeifenbauform. JAKOB ADLUNG (1699–1762) beschreibt den Sachverhalt knapp und doch ebenso umfassend wie treffend:

»Vox humana, die Menschenstimme, ist ein Schnarrwerk [Lingualregister] 8', weil ein Mensch deutlich so tief singen kann. [...] Wie aber die Körper [Resonatoren] eigentlich zu machen, sammt den Mundstücken [Kehlen], daß es einen Klang von sich gebe, welcher der Menschenstimme gleicht, ist eine hohe Frage. Nichts ist schöner, als die Stimme des Menschen: nichts ist aber auch schwerer nachzumachen, als eben dieselbe. Deswegen hat man zwar vielerley Inventionen; aber keine hat vollkommen geleistet, was sie [hat] leisten sollen, und es wird doch allemal ein großer Unterschied bleiben unter [zwischen] dem Klang einer solchen Pfeife, und der natürlichen Stimme eines Menschen. Es thut auch etwas, daß die Orgelmacher in der Anatomie nicht bewandert sind, und daher nicht wissen, wie der Klang in unsere [sic!] Kehle formieret wird.« [Adlung 1768, 155 f.]

---

<sup>7</sup>Vgl. die Hörbeispiele im Appendix (S. LXXI).



**Abbildung 2.1:** Schematische Darstellung einer klassischen französischen *Vox humana* des 18. Jahrhunderts [nach Bedos 1977, Pl. XVIII, Fig. 141]

Und an anderer Stelle noch knapper:

»Vox humana, die Menschenstimme, ist desto schwerer nach zu machen, weil die Arbeiter sich bis ietzo nicht vereinigen [einig werden] können, ob sie in ihrer Vollkommenheit solle eine weibliche, oder männliche Stimme vorstellen, einen Discant [Sopran], Alt, Tenor oder Baß.« [Adlung 1758, 476 f.]

FRANÇOIS L. DOM BEDOS DE CELLES (1709–1779), Zeitgenosse Adlungs und profunder Kenner insbesondere des französischen Orgelbaues, stellt in seinem 1766–1787 im Auftrag der Pariser Akademie der Wissenschaften erschienenen Grundlagenwerk »*L'art du facteur d'orgues*« lapidar fest:

»Jeder gestaltet sie [die Pfeifen der VH] nach seiner Veranlagung; doch besser gelingen sie dadurch nicht. [...] Seine Güte [des Registers *Vox humana*] besteht darin, die menschliche Stimme so vollkommen als möglich nachzuahmen. Dies gelingt selten.« [Bedos 1977, 71 f.]

DIDEROTS »*Encyclopédie*« geht recht detailliert auf die Konstruktionsmerkmale der VH ein (vgl. Abb. 2.1<sup>8</sup>):

»Voix humaine, jeu d'orgue, ainsi nommé, parce qu'il imite assez bien, quand le jeu est bien fait, la voix de l'homme, est un jeu de la classe des jeux d'anches: il est d'étain, & sonne l'unisson de la trompette, aux anches de laquelle les anches sont égales; mais son corps qui est de plus grosse taille, & n'a que le quart de longueur. (Voy. la fig. 40. Pl. d'orgue a b,)<sup>9</sup> est le corps du tuyau qui est à moitié fermé par le haut avec une plaque d'étain a, dont la forme est un demi-cercle.« [Diderot 1765, 438]<sup>10</sup>

<sup>8</sup>Die Abbildung der *Vox humana* in der »*Encyclopédie*« war aufgrund schlechter Druckqualität der verfügbaren Vorlagen nicht verwendbar. Daher wurde auf die sehr ähnliche Abbildung bei [Dom Bedos 1977] zurückgegriffen.

<sup>9</sup>Recte: Recueil des Planches, Tome 4 (1767); »Lutherie«, Pl. IX, Fig. 48

<sup>10</sup>Dasselbe Verfahren beschreibt auch DOM BEDOS, um »die natürliche Stimme des Menschen zu imitieren« [Bedos 1977, 316].

»Vox humana, ein Orgelregister, das so genannt wird, weil es, wenn es gut gemacht ist, die Stimme des Menschen ganz gut imitiert, ist ein Register aus der Klasse der Zungenregister: Es besteht aus Zinn und klingt unisono mit der Trompete [es klingt in der selben Tonhöhe wie die *Trompette*, die stets im 8'-Ton steht], da seine Zungen[blätter] dieselbe Länge haben; aber die Körper [Resonatoren], die erst in ihrer Mitte ihren größten Durchmesser erreichen, sind nur ein Viertel so lang [wie die des Registers *Trompette*] (siehe Fig. 40 auf der Bildtafel zur Orgel a,b). ist der Resonator, der an seinem oberen Ende zur Hälfte mit einer Zinn-Platte in Form eines Halbkreises verschlossen ist.«

Synonym zur lateinischen Bezeichnung des *Registers* wurde insbesondere im 17. Jahrhundert auch die griechische Form *Anthropoglossa* verwendet. Dies geschah insbesondere dann, wenn es galt, zwischen dieser als edler angesehenen Bauform der *Regal*-Familie und den als *Zooglossa* bezeichneten gewöhnlichen Bauformen zu unterscheiden. So heißt es bei ATHANASIUS KIRCHER:

»Est aliud adhuc fistularum genus, quod nos nō incongruè Anthropoglosson appellamus, propè verum enim sermonem humanum risumque mentitur, adeoq; miram cōfert gratiam, vt in toto organico systemate nihil admiratione dignius, quodque animum auditorium vehementius rapiat, inueniatur. Estq. fistula mixta ex Zooglossa & tubo non purè cylindraceo sed conico cylindraceo cui Zooglossa inseritur [...].« [Kircher 1650, 514]

»Es gibt außerdem [außer der Zooglossa] noch eine andere Art von Pfeifen, die wir nicht unpassend Anthropoglossa nennen, denn nahezu wirklichkeitsgetreu täuscht sie menschliche Rede und menschliches Gelächter vor, und bereitet so sehr wunderbare Annehmlichkeit, dass in der ganzen Orgel nichts der Bewunderung Würdigeres gefunden werden kann und nichts, was das Herz und den Geist der Zuhörer heftiger erfasste. Und zwar ist diese Pfeife zusammengesetzt aus der Zooglossa und einem nicht rein zylindrischen, sondern konisch-zylindrischen Resonator, an den die Zooglossa angefügt wird [...].« (Übersetzung Dr. ELISABETH LINK)

Die Register *Cromorne* und VH unterschieden sich folglich nur durch die Form ihres *Resonators*. KIRCHER gibt abschließend in drei »Regulæ« höchst detaillierte Anweisungen, wie man die fraglichen Orgelpfeifen bauen müsse, damit sie den beschriebenen Klangeindruck hervorrufen.

Allerdings wurden zumindest vereinzelt auch andere *nicht-langbecherige Lingualregister* mit der menschlichen Stimme verglichen. So heißt es in einem Entwurf für den Umbau der Orgel der Lübecker Petrikirche von 1586 über das *Register Dulzian* (»tusan«):

»Eyn tusan, eyn stemme, also genennet, [weil] seine leylichheyt zu verwonderen is un anders nicht mach vorgleichen werden als eyn naturliche menschenstem.« [Eberlein 2009, 169]

Auch noch zu Beginn des 20. Jahrhunderts sieht man in der VH ein ernstzunehmendes Problem. Der Schweizer Organologe CARL LOCHER (1843–1915) führt aus:

»Vox humana ist ein von fast jedem Orgelbauer anders konstruiertes 8<sup>e</sup> Zungenregister, welches die Stimme nachahmen soll. Trotz der vorzüglichsten Konstruktion bleibt aber (etwa mit Ausnahme der mittleren Lage) immer ein etwas näselnder Metallton bemerkbar. Der Schall der Zungenpfeifen entsteht nach Helmholtz durch die intermittierenden Luftstöße, welche durch die Schwingungen hindurchbrechen. [...] Hartes, unnachgiebiges Material, wie das der Messingzungen [Zungenblätter für Lingualpfeifen werden in aller Regel aus Messing gefertigt], wird die Luftstöße viel mehr abgerissen hervortreten lassen als weiches, nachgiebiges. [...] Hierin haben wir augenscheinlich den Grund zu suchen, warum die menschlichen Gesangstöne gut gebildeter Kehlen sich durch die Weichheit von allen Klängen der Zungenpfeife, also auch der best gebauten Vox humana, wesentlich unterscheiden und auszeichnen werden. [...] Verbessert wird die Vox humana durch eigentlich ausserhalb ihr selbst liegende lokale Eigentümlichkeiten, wie z.B. Anbringung in einer abgelegenen Tonhalle<sup>11</sup> [...], geschmackvolle Verwendung des Tremulant oder Crescendotrittes, etc.« [Locher 1904, 125 f.]

Der Musikwissenschaftler ROLAND EBERLEIN (\*1959) skizziert in seinen umfassenden Kompendien zur Geschichte der Orgel(register) die Entstehungsgeschichte der VH aus historisch-kritischer Sicht schliesslich wie folgt:

»In Frankreich wurde spätestens seit ca. 1550 das mit gezogenem Tremulanten gespielte Regal als Voix humaine bezeichnet. Der Name der Registrierung ging bald [als pars pro toto] auf das Regal über. [...] Alle diese [Bau-]Formen galten schon im 18. und 19. Jahrhundert als klanglich weit entfernt von der Menschenstimme.« [Eberlein 2007b, 817]

---

<sup>11</sup>Gemeint ist die Aufstellung räumlich getrennt von der eigentlichen Orgel in einem separaten, besonders schallgedämmten *Schwellwerk*.

»Von ca. 1550 an wurde das Regal in den Niederlanden vereinzelt als *Vox humana*, Stimmekens, Kindstem oder - bei flämisch beeinflussten Orgelbauern in Frankreich - *jeu d'enfants* genannt.« [Eberlein 2009, 701]

CHRISTHARD MAHRENHOLZ (1900–1980), Theologe und Pionier der »Orgelbewegung«, erklärte beinahe ein Jahrhundert zuvor die Benennung des *Registers* hingegen mit der Musizierpraxis des Barock, betont aber auch, dass es sich bei der Bezeichnung VH eher um einen Gebrauchshinweis handelte denn um eine Charakterisierung der Klangeigenschaften:

»Sie [die Lingualregister der Regalbauweise] entsprachen den konzertierenden Solostimmen neben den instrumentalen Chören in der Barockmusik. Hier erklärt sich dann, warum die kurzbecherigen Regalarten aller Formen auch mit dem Namen *Vox humana* belegt werden, ein Ausdruck, den Prätorius<sup>12</sup> zwar nicht nennt, der sich aber schon 1598 in Lüdingworth und anderswo findet und wohl kaum aus Frankreich importiert ist. Die *Vox humana* will die Menschenstimme nicht nachahmen, wie die italienischen *Voci umane* [siehe Glossar], sondern der Ausdruck besagt, dass das betreffende Orgelregister eine Funktion ausübt, die sonst in der Barockmusik der konzertierenden Menschenstimme zufällt. Daher kommt es auch, dass der Name *Vox humana* mit keinem bestimmten Register verbunden ist.«<sup>13</sup> [Mahrenholz 1928, 27]

Späterhin revidierte MAHRENHOLZ seine frühere Auffassung: Die VH oder »Anthropoglossa« genannten Bauform ist stets mit für Regale verhältnismäßig langen Resonatoren ausgestattet. Die urwüchsigeren *Regale* mit sehr kurzen Resonatoren werden bisweilen kontrastierend »Zooglossa« genannt, weil ihr Klang an das Blöken von Kälbern erinnere [Mahrenholz 1982, 166].

In der Retrospektive eröffnen sich noch weitere Deutungsmöglichkeiten des Phänomens VH. So sieht FRANK-HARALD GRESS einen historischen Zusammenhang zwischen Musikphilosophie und dem Registerbestand der zeitgenössischen Orgeln:

»Die Musikästhetik der Aufklärungsepoche ist durch die Nachahmungstheorie gekennzeichnet, die sich ab 1700 entwi-

<sup>12</sup>MICHAEL PRAETORIUS (1571(?)-1621), deutscher Komponist, Organist und Musiktheoretiker. Autor des »*Syntagma musicum*« (1615/19), auf das M. hier Bezug nimmt.

<sup>13</sup>Gemeinsam ist allerdings allen VH-Formen der für *Regale* verhältnismäßig lange und teilgedeckte *Resonator*.



ckelte. Ihr Stichwort hieß: Nachahmung der Natur. Klangimitationen und Klanganalogien fanden in dieser Zeit verstärktes Interesse. Die anlässlich der Einweihung neuer Orgeln verfaßten zahlreichen ›Carmina‹ schenken den imitatorischen Elementen bezeichnenderweise größeres Interesse als den orgeleigenen Klangmitteln. Vor allem die ›Vox humana‹ wurde immer wieder freudig bestaunt. ›Kann was natürlicher, als Vox humana, klingen. . . ?‹, schrieb 1733 Wilhelm Friedemann Bach [1710–1784] über die Orgel der Dresdner Frauenkirche [Gottfried Silbermann 1736 (!)]. Johann Ludwig Krebs [1713–1780] stellte 1737 über die Orgel in Ponitz [G. Silbermann 1737] fest: ›Da hörte man den Klang von rechten Menschen=Stimmen. . . ‹, und Jakob Adlung [1699–1762] berichtet über die Vox humana der Greizer Orgel: ›Ihr soll nichts fehlen als die Worte‹.« [Gress 2007, 27]

Dieser Ansatz erklärt indessen zwar das Verständnis der VH im 18. Jahrhundert, nicht jedoch ihre Entstehung und Namensgebung zweihundert Jahre zuvor.

Der Journalist, Dichter und Organist CHRISTIAN F. D. SCHUBART (1739 – 1791) charakterisiert die VH der Orgel in der Mannheimer Konkordienkirche aus der Werkstatt der Gebrüder STUMM (1765) so:

»Nirgends fand ich die Menschenstimme täuschender und reiner als hier auf dieser Orgel.«<sup>14</sup> [Bösken 1981, 125]

Der Musikhistoriker CHARLES BURNEY (1726–1814) thematisiert in seinen Berichten über das Musikleben auf dem europäischen Festland pointiert den Widerspruch zwischen Anspruch und Wirklichkeit bei der VH:

»[In der Orgel der Dominikanerkirche in Frankfurt/Main war] auch die Vox humana nicht sonderlich angenehm, oder der Menschenstimme ähnlich, ob man hier gleich viel Wesens daraus machte.« [Burney 1773, 61]

Anders als alle anderen *kurzbecherigen Lingualregister* fand die VH auch im romantisch geprägten 19. Jahrhundert noch Verwendung in neugebauten Orgeln [Eberlein 2009, 702]. Dass die VH noch bis ins 20. Jahrhundert hinein einen gewissen Sonderstatus gegenüber den sonstigen Orgelregistern besaß, zeigt ein Bericht aus dem Jahre 1905 über die soeben eingeweihte Orgel des Berliner Doms (die damals größte Orgel Deutschlands). Gleich

<sup>14</sup>Diese nicht besonders große Orgel verfügte im Übrigen über gleich zwei verschiedene Vox humana-Register [Eberlein 2011, 200].

zu Beginn des Artikels werden aus insgesamt 113 *Registern* einer nach damaligen wie heutigen Urteilen exzeptionellen Orgel zwei *Register* als besonders herausgegriffen: Das idiophone *Glockenspiel* als absolute Rarität in einer Orgel und die VH [Anonymus 1905, 508].

Insbesondere die hier eingeflochtenen historischen Zitate von BACH und KREBS – beide Organisten und Komponisten von Rang – belegen eindrücklich, dass man seinerzeit zumindest manchen VH-*Registern* eine nicht unerhebliche Authentizität zuschrieb. Zunächst offen bleiben muss in diesem Zusammenhang die Frage nach dem Grund hierfür. Nehmen wir heute diese Äußerungen viel ernster und wörtlicher, als sie seinerzeit gemeint waren? War die Wahrnehmung eine vollständig andere als heute, nahm man damals also Klänge, denen wir heute höchstens vage Ähnlichkeit attestieren würden, als realistisch wahr? Oder sollten alle VH, die seinerzeit so menschenähnlich klangen, verloren gegangen sein und nur die weniger guten die Zeiten überdauert haben? Eine wenig wahrscheinliche Annahme. Plausibel erscheint hingegen eine Kombination aus den ersten beiden Vermutungen, also eine nicht wortwörtlich zu nehmende, blumige Sprache gepaart mit grundlegend anderen Wahrnehmungskriterien.

Eine weitere mögliche Erklärung für den hier gezeigten Widerspruch zwischen der VH und der menschlichen Stimme könnte das Stimmideal sein, das sich im 18. Jahrhundert erheblich wandelte. Im Zuge der Recherchen stieß der Autor im Internet auf Zeugnisse einer heute noch auf Sardinien geübten Tradition des Volksliedes, des »Canto a tenore«. <sup>15</sup> Hierbei handelt es sich um einen von i. d. R. vier Männern a capella in der Technik des Obertonsingens ausgeführten mehrstimmigen Gesang. Den »Oche« genannten, einen Text singenden Vorsänger begleiten drei als »Bassu« (Bass), »Contra« (Bariton) und »Mesu oche« (Alt) bezeichnete Sänger mit auf Nonsenssilben gesungenen Akkorden. Der hierbei angewendete Kehlgesang weist eine frappante Ähnlichkeit mit dem Klang barocker VH auf. <sup>16</sup>

Man könnte nun vermuten, dass es ähnlich urwüchsige Traditionen des Volksliedes auch in anderen Gegenden Europas gab, die erst durch das aus der Kunstmusik stammende und sich im 18. und besonders 19. Jahrhundert verbreitende Ideal des »Belcanto« verdrängt wurden. Diese Theorie lässt sich indes quellenmäßig nicht belegen, wie dem Autor von mehreren hierzu befragten Fachleuten mitgeteilt wurde. Grund hierfür ist, dass die volkstümlichen Traditionen des Musizierens erst sehr spät als Gegenstand wissenschaftlicher Untersuchungen erkannt wurden.

<sup>15</sup>Den initialen Hinweis hierzu verdanke ich Kantor JOACHIM SCHNEIDER, Simmern.

<sup>16</sup>Vgl. bspw. das Hörbeispiel auf der CD-R im Appendix bzw. unter <https://www.youtube.com/watch?v=rQuVQ1dkenE> (zuletzt besucht am 07.06.2014)

»[...] Das Problem ist in der Tat der Quellen-Nachweis. Wie Sie richtig vermuten, müssen wir davon ausgehen, dass in Europa zumindest vor 1700 ganz anders gesungen wurde, als wir das heute für den Kunstgesang gewohnt sind. Diese Tradition ist relativ jung und stammt weitestgehend vom Belcanto-Ideal des 19. Jahrhunderts.

Wie genau die Stimmen allerdings damals geklungen haben, ist schwer auszumachen, das ist eigentlich auch die Crux der historischen Aufführungspraxis.

Die heute noch lebendigen Traditionen auf Sardinien und Corsica haben ja einige Mittelalter-Ensembles inspiriert, so zu singen (besonders das Ensemble »Organum« unter Marcel Peres). Da ist sicher viel Wahres dabei. [...]

Es gibt vereinzelte Hinweise über sehr tiefe Stimmen (bassi profondi), z. B. bei Praetorius im 16. Jh. über die Bayerische Hofkapelle, die wohl auch aus einer Volkstradition kamen.

Vereinzelte finden wir in den Reiseberichten etwa von Charles Burney [vgl. Burney 1722, 222 u. 271] Hinweise, dass man in Italien auf Volkssänger gestoßen ist, die sehr kehlig und mit Bruststimme und in Quinten gesungen haben.

Speziell für Frankreich wüsste ich momentan keine Quelle – es ist aber fast sicher anzunehmen, dass neben der Kunstmusik auch dort solche volkstümlichen Traditionen existiert haben.« (persönliche Mitteilung von Prof. Stefan Morent, Tübingen, vom 14.05.2013)

Einen Hinweis darauf, dass die Gesangskultur im westeuropäischen Mittelalter zumindest zum Teil derjenigen auf Sardinien geähnelt haben dürfte, bringt beispielsweise MÜLLER-HEUSER [1997, 85 f.]. Ähnliche Indizien für ein eher »kehlig« geprägtes Gesangsideal im nördlichen Teil Europas finden sich bei THORNTON [1980, 148 f.].

Sollte die Annahme stimmen, dass auch im Frankreich des 15. und 16. Jahrhunderts, also zur Ursprungszeit der VH, eine vergleichbare volkstümliche Singetradition existiert hat, wäre das Rätsel über die angebliche Stimmähnlichkeit der VH gelöst. Spätere Generationen wussten nicht mehr um den ursprünglichen Bezug und maßen die »Stimmqualität« der VH an den Idealen des Belcanto, was freilich nicht funktionieren konnte.

Die Obsession von der theoretischen Möglichkeit einer möglichst großen »Authentizität« der VH ließ vereinzelte die Fantasie ausufern, wie die Sage über das *Register* der großen Orgel der ehemaligen schwäbischen Benediktiner-Abtei Weingarten (JOSEPH GABLER 1750) zeigen mag:

»Gabler arbeitete jahrelang an der Lösung des Problems [einer möglichst naturgetreuen Vox humana]; trotz aller Auswahl der Holzarten und Metallmischungen kam er nicht zu dem gewünschten Resultat. [...] Gabler verschrieb mit Blut dem Teufel seine Seele und bekam dafür ein Stück Metall, das er zum Pfeifenguß verwerten sollte. – Das teuflische Metall wird in die Pfeifen gegossen [die Pfeifen werden aus dem Metall gegossen] und herrlich erklang die Vox humana, wie die menschliche Stimme. Aber, o weh! Anstatt heiliger Melodien sang sie von der Lust der Welt, so daß viele Mönche ihre stille Klausur verließen und sich in die Weltfreuden stürzten. [...] Gabler wurde der Prozeß gemacht und er sollte mitsamt dem unheimlichen Register im Klosterhof verbrannt werden. Zuvor aber sollte er wenigstens annähernd Ersatz schaffen. Dies gelang dem Meister nun so gut, daß ihm der Abt gnädig das Leben schenkte.« [Bärnwick 1922, 49 f.]

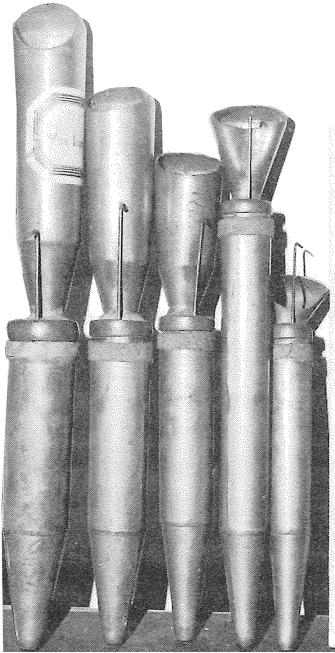
Einen wahren Kern besitzt diese abenteuerliche Legende jedoch: Bei der Formgebung der Resonatoren seiner *VH*-Pfeifen in Ochsenhausen (1734/55) und Weingarten ließ sich GABLER unverkennbar durch die äußere Form des menschlichen Kehlkopfes inspirieren, was vermutlich einiges zur »Sagenhaftigkeit« dieser Pfeifen beigetragen hat (Abb. 2.2).

WERNER LOTTERMOSER analysierte GABLERS Konstruktionsvariante der *VH* hinsichtlich der spektralen Zusammensetzung ihres Klanges und attestiert ihr eine deutliche Nähe zur menschlichen Stimme:

»[Das Maximum der Abstrahlung fällt auf den 7. Teilton.] Darauf folgte ein zweites Frequenzgebiet mit dem 17. Teilton im Maximum sowie höhere Teiltöne. Diese Struktur entspricht etwa der menschlichen Stimme, so daß auch der Klang ein ähnliches Timbre hatte.« [Lottermoser 1983a, 134 f.]

Vermessen wurde hierbei jedoch der Ton C (- 70 Hz), der außerhalb des Stimmambitus beinahe eines jeden Menschen liegt und sich dementsprechend in keiner typischen Gesangslage befindet.

Es bleibt an dieser Stelle abschließend festzuhalten, dass es sich bei der *VH* ursprünglich gar nicht um die Bezeichnung einer konkreten Pfeifenbauform handelte und es auch nicht den Anspruch gab, dem Klang der menschlichen Sprachstimme so nahe als möglich zu kommen. Erst späterhin entstand durch Unkenntnis der Herkunft des Begriffes der Glaube, Name und Wirkung müssten in einem konkreten, realistischen Zusammenhang stehen.

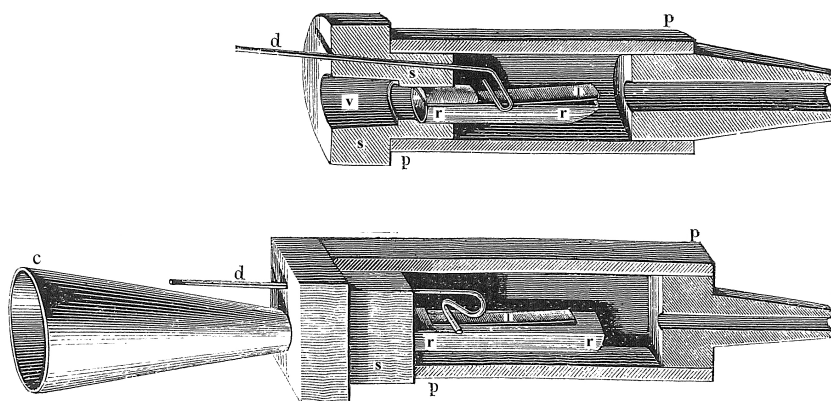


**Abbildung 2.2:** Pfeifen von GABLERS Vox humana in Weingarten (Töne  $C$ ,  $c^0$ ,  $c^1$ ,  $c^2$  und  $c^3$ ). [Lottermoser 1983a, 135]

## 2.2 Allgemeine Konstruktionsmerkmale der »Vox humana«<sup>17</sup>

Das Funktionsprinzip einer *Lingual-* oder auch *Zungenpfeife* ist dem eines Holzblasinstruments mit einfachem *Rohrblatt* – beispielsweise einer Klarinette – sehr ähnlich: In Abb. 2. 3 oben dargestellt sieht man ein dünnes, meist schmalrechteckiges Blättchen aus Metall (l) – *Zunge* oder *Zungenblatt* genannt – das auf einem Rohrsegment aus Holz oder Metall, der *Kehle* (r) liegt. Dieses ähnelt in seiner Form einem der Länge nach halbierten hohlen Zylinder mit geschlossenem unteren Ende. Die *Kehle* mit der darauf liegenden *Zunge* ist mit ihrem offenen oberen Ende in der mit-tigen Bohrung der Unterseite eines eckigen oder runden Blocks aus Holz

<sup>17</sup>Die folgenden Ausführungen sind modifiziert der Masterarbeit des Autors entnommen (vgl. [Brackhane 2011, 145-148]). Für *orgelbauliche Fachbegriffe* vgl. das Glossar im Anhang). Die in Klammern angegebenen Buchstaben beziehen sich auf Abb. 2.3 (S. 20).



**Abbildung 2.3:** *Lingualpfeifen mit aufschlagendem (oben) und durchschlagendem Zungenblatt (unten) [nach Helmholtz 1863, 155].*

oder Blei, dem *Kopf* (s) befestigt. Auf der Oberseite des Kopfs geht diese Bohrung (v) in einen *Resonator*, den sog. *Becher* (c) über. Die Unterseite des *Kopfes* mit *Kehle* und *Zunge* steckt in einer zylindrisch-konischen oder rechteckigen Röhre, dem *Stiefel* (p), an dessen Unterseite sich ein Loch für die Luftzufuhr – man spricht im Orgelbau vom *Wind* – befindet. Strömt nun durch dieses *Fußloch* Wind in den *Stiefel*, entsteht in diesem ein Überdruck. Hierdurch wird die auf der *Kehle* liegende *Zunge* hochgedrückt, der Wind strömt durch *Kehle*, *Kopf* und *Becher* ins Freie. Nach dem *Bernoulli-Prinzip* entsteht nun aber im *Stiefel* ein Unterdruck, der die *Zunge* wieder fest auf der *Kehle* aufliegen lässt und den Luftstrom stoppt, wodurch ein neuerlicher Überdruck im *Stiefel* entsteht. Die Federkraft des *Zungenblattes* unterstützt diesen Effekt zusätzlich. Das Prinzip der Luftanregung ist also im Falle der *Lingualpfeife* ein ähnliches wie das der menschlichen Stimmgebung, wenngleich im letzteren Falle zwei Elemente (die beiden Stimmlippen) gegeneinander schwingen, wogegen in ersterem nur ein Element – die *Zunge* – schwingt und das zweite – die *Kehle* – starr ist.

Der Klangcharakter einer *Zungenpfeife* ist von zahlreichen Faktoren abhängig, so beispielsweise von der Breite und Stärke eines *Zungenblattes* (l) in Relation zu seiner Länge. Die Konstruktion der *Kehle* (r) ist ebenfalls maßgeblich: Sind die Seiten der *Kehle* parallel und der vordere Abschluss gerundet, bewirkt dies u.a. ein schnelleres Ansprechen des *Zungenblattes* auch bei niedrigem Druck. Eine nach diesen Grundsätzen

konstruierte *Kehle*, eine sog. *Französische Löffelkehle*, benutzte beispielsweise auch WOLFGANG VON KEMPELEN, für seine endgültige Form der Sprachmaschine (vgl. Abb. 8.2 auf S. 171). Der andere Typus von *Kehlen* wird häufig bei eher schmalen *Zungenblättern* angewendet; die Kehlen-seiten laufen hier nicht parallel, sondern leicht konisch, die Oberseite der *Kehle* ist nicht wie bei den *französischen Kehlen* vollständig offen, sondern teilweise geschlossen. Der Klangcharakter ist hierdurch weniger kraftvoll.

Nicht zuletzt sind Form und *Mensur* des auf dem *Kopf* aufsitzenden *Resonators* (c), des sog. *Schallbechers*, von ganz erheblicher Bedeutung für den Klang. Grundsätzlich gilt, dass konische (also vom *Kopf* an nach oben hin immer weiter werdende) *Schallbecher* sämtliche Teiltöne der *Zunge* gleichmäßig verstärken, während zylindrische *Resonatoren* dies nur bei den ungeraden Teiltönen tun und damit einen helleren, »leichteren« Klang erzeugen. Je kürzer der *Resonator* ist, desto weniger formenden Einfluss besitzt er natürlicherweise auf den Klang.

Bei den meisten *Lingualregistern* ist die Länge eines *Resonators* so bemessen, dass seine Eigenresonanz mit der Tonhöhe (=Schwingungsfrequenz) der *Zunge* übereinstimmt. Sind die *Resonatoren* erheblich kürzer (wie bei der VH), bewirkt dies einen grundtonärmeren, schnarrenderen Klang.

Es sei indessen nochmals betont, dass gerade bei der VH hinsichtlich der Bauform eine nahezu unüberschaubare Vielfalt herrscht, insbesondere was die Form der *Resonatoren* betrifft. Diese können schlichte kurze Trichter sein, genauso gut aber auch recht komplexe zusammengesetzte geometrische Figuren, die aber gleichwohl immer innerhalb der Gattung der *kurzbecherigen Zungenstimmen* bleiben. Eine generalisierende Aussage zur Bauform kann hier noch nicht einmal für eine Epoche innerhalb einer Region getroffen werden, da jeder Orgelbauer an diesem *Register* seinen ganzen Einfallsreichtum erproben konnte und wollte. Generell besitzen VH-Pfeifen aber für Pfeifen der *Regal*-Bauweise verhältnismäßig langen *Resonatoren*, die zudem noch sehr häufig an ihrer Mündung teilweise verschlossen werden, um eine Dämpfung des Klanges zu erreichen.

Der Sprachsynthesepionier CHRISTIAN GOTTLIEB KRATZENSTEIN (1723-1795, vgl. auch Kapitel 3.2) änderte das generelle Funktionsprinzip einer *Lingualpfeife* für seine Zwecke in einem entscheidenden Punkt ab: Er konstruierte *Zunge* (l) und *Kehle* (r) so, dass die *Zunge* nicht mehr auf dem Rand der *Kehle* aufliegt (»*aufschlägt*«, Abb. 2.3 oben), sondern ganz exakt in die Öffnung der *Kehle* hineinpasst (»*durchschlägt*«) und sie auf diese Art in ihrer Ruhestellung verschließt (Abb. 2.3 unten). Strömt nun Luft an die *Zunge*, so schwingt sie nicht nur hoch und wird auf dem Rückweg durch die Ränder der *Kehle* aufgehalten, sondern kann in glei-

chem Maße, wie sie zuvor hoch geschwungen ist, nun auch nach unten schwingen.

KRATZENSTEIN hatte damit das Prinzip der *durchschlagenden Zunge* (im Gegensatz zur *aufschlagenden Zunge*) entdeckt. Zwar war diese Konstruktion in China bereits seit vielen Jahrhunderten als Eigenschaft der *Mundorgeln* bekannt, doch war sie bis dato in Europa noch nicht als neuartiges Anregungskonzept erkannt worden. Im Orgelbau insbesondere des 19. Jahrhunderts sollte diese Entdeckung ganz wesentliche Bedeutung erhalten [Ahrens 2002, Braasch 2004]. Zunächst wurde diese Modifikation der VH jedoch durch KRATZENSTEIN zum Gegenstand der seinerzeit noch in den Kinderschuhen steckenden Sprachsyntheseforschung (vgl. Kap. 3.2.3 und 7).

Allerdings griffen sowohl KRATZENSTEIN als auch KEMPELEN faktisch nur auf das grundsätzliche Anregungskonzept der Lingualpfeifen zurück, das eigentliche Charakteristikum der VH, der kurze, teilgedeckte *Resonator*, fand dabei keine Berücksichtigung.

## 2.3 Ein Prototyp der Sprachsynthese?

Die zwei so verschieden erscheinenden Bereiche Orgelbau und Sprachsynthese stehen bzw. standen lange Zeit in erstaunlich intensiver Beziehung zueinander. Entscheidend hierfür sind zwei Einzelaspekte, die jedoch eng miteinander verknüpft sind: Einerseits beobachtete man früh, dass die in Orgeln verwendeten *Zungenpfeifen* und hier insbesondere die *Regale* einen stark von formantartigen Strukturen geprägten Klang erzeugen, der bisweilen besonders in tieferen Lagen an gesungene (statische) Vokale erinnert. Durch diese Feststellung wurden die oben dargelegten Missverständnisse in Sachen VH noch verstärkt [Lottermoser 1983a, 59 und 62]. So berichtete beispielsweise LEONHARD EULER (1707–1783) über seine Beobachtung, dass die »Vokalqualität« bei VH-Pfeifen von deren Tonhöhe und den Proportionen des *Resonators* abhängt [Euler 1862, 798 f.] (vgl. auch Kapitel 3.2.2).

Hinzu kam zudem die Vorstellung, dass die Orgel mit ihrer Möglichkeit, Tonhöhe und Klangfarbe über verschiedene Pfeifenbauformen nahezu unbegrenzt variieren zu können, als Vorbild für eine maschinelle Sprachsynthese dienen könne. In der zeitgenössischen Literatur finden sich mehrere Belege für diese Vorstellung, eine Sprachsynthese könne funktionieren wie eine Orgel (vgl. [Euler 1775, 246 f.] sowie Kap. 3.2.2). Schon im 17. Jahrhundert war diese Vorstellung, wie am Ende des Kapitels 1 gezeigt, offenbar so weit verbreitet, dass BERNARD LAMY mit seinem Gedanken-



experiment ein für alle Mal die praktische Unmöglichkeit dieser Theorie zu zeigen versuchte.

Einen konkreten physiologischen Bezug stellt der französische Arzt und Botaniker DENIS DODART (1634-1707) her, indem er die (absolute) Länge einer VH-Pfeife in Beziehung zur Länge des menschlichen Ansatzrohrs setzte, die in etwa gleich sei [Dodart 1700, 249].<sup>18</sup> In seiner Argumentation ergibt sich aus dem Dreiklang von Namen, Funktionsweise und Klang ein so starker Bezug zur menschlichen Stimme, dass die Folgerung, es handele sich bei der VH um den Prototyp der Sprachsynthese schlechthin, geradezu zwingend erscheint.<sup>19</sup>

Wie bereits dargestellt, war es mit der VH eine bestimmte Unterform der *Regale*, der das besondere Augenmerk der Sprachsyntheseforscher galt. So schreibt HINDENBURG in seinem Bericht über die Sprachmaschine KEMPELENS:

»Die so genannte *vox humana* der Orgel, so weit sie auch von der Menschenstimme noch [!] entfernt ist, kommt ihr gleichwohl viel näher [als die Synthese der Sprachmaschine].«  
[Hindenburg 1784, 54]

Dass auch ein halbes Jahrhundert nach KEMPELEN und insbesondere KRATZENSTEIN die Frage nach einer ihrem Namen entsprechenden Funktionalität der VH sich nicht nur auf musikalische Kreise beschränkte, zeigen die Einlassungen des englischen Gelehrten und Ingenieurs ROBERT WILLIS (1800–1875):

»Die zweite Verbindung von Röhren [Resonatoren] und Zungen findet sich bei der *Vox humana*, wo sehr kurze Röhren, die nicht mit der Zunge im Einklang stehen [deren Eigenresonanz nicht mit der Schwingungsfrequenz der Zunge übereinstimmt], angewandt werden. Über die Principien dieser Zusammensetzung [dieser Bauform von Lingualregistern], die den Orgelbauern immer ein Räthsel war, ist man bis jetzt noch nicht im Klaren.« [Willis 1832, 431]

Über den Zusammenhang von Tonhöhe und »Vokalqualität« bei VH-Registern führt WILLIS anschließend aus:

<sup>18</sup>Vgl. hierzu die Gegenüberstellung der Längenmaße einer klassischen VH und des menschlichen Vokaltrakts auf S. 157!

<sup>19</sup>Zu Dodarts Entlastung sei angemerkt, dass er in Hinblick auf die physiologischen Aspekte der Sprachproduktion vielen seiner Zeitgenossen weit voraus war (vgl. [Dodart 1700, 255 ff.]).

»Versteht man unter einem »Vox-humana-Register« ein solches, welches in allen seinen Tönen denselben Vocallaut giebt, so ist die Sache leicht durch die von mir auseinandergesetzten Principien abgethan. Man muss sich indess erinnern, dass es für solch ein Register immer eine natürliche Gränze in der Höhe giebt, da es unmöglich ist, einen und denselben Vocallaut auf einen höheren Ton, als den ihm eigenthümlichen zu übertragen. So könnte ein O-Register nicht weiter als bis  $c^{II}$ , ein A-Register nicht weiter als bis  $f^{III}$ , und ein E-Register nicht weiter als bis  $d^{IV}$  ausgedehnt werden.«<sup>20</sup> [Willis 1832, 432]

Die VH besitzt also keinerlei verlässliche Konstruktionsformen oder detailliert konstante Klangeigenschaften. Der Grund für die Wahl ausgerechnet dieses Orgelregisters als Anregungsgrundlage für eine mechanische Sprachsynthese scheint zumindest auch in den beiden bereits zuvor erläuterten grundlegenden Missverständnissen zu bestehen: Zum einen war das Wissen über die Etymologie des Registernamens offenbar nicht mehr präsent, so dass man anscheinend davon ausging, es müsse einstmals *VH-Register* gegeben haben, die erheblich »menschlicher« klangen. Den zweiten fundamentalen Irrtum lässt LEONHARD EULER in seinen »*Briefen an eine deutsche Prinzessin*« deutlich werden, nämlich den Glauben daran, im Instrument Orgel quasi eine Vorform der Sprachsynthese zu haben:

»In vielen Orgeln findet man ein Register, das Vox humana (die Menschenstimme) genannt wird; gemeiniglich aber macht sie nur Töne, die den Vocal ai [i] oder ae [e] nachahmen.<sup>21</sup> Ich zweifle nicht, daß man mit einigen Veränderungen auch die übrigen Vocalen a, e, i, o, u würde herausbringen können; aber alles dieses würde noch nicht hinreichen, ein einziges Wort der menschlichen Stimme nachzumachen; denn wie wollte man die Consonanten mit ihnen verbinden, die so viele Modificationen der Vocalen sind? Unser Mund ist so bewundernswürdig eingerichtet, daß es uns unmöglich ist, den Mechanismus, der zu diesem so gemeinen Gebrauche desselben gehört, zu ergründen. [...] Ein großer Beweiß von dem wunderbaren Baue unsers

<sup>20</sup>Bei der Bezeichnung der Tonhöhen besteht Verwechslungsgefahr: Die von Willis im englischen Originaltext verwendete Bezeichnungsweise ist nicht die übliche musikalische Notation, sondern bezieht sich als wissenschaftliche Notation höchstwahrscheinlich auf das Kontra-C (32 Hz) als Basis. Der von Willis als  $c^{II}$  bezeichnete Ton entspräche demnach nicht dem Ton  $c^2$  nach musikalischer Notation (532 Hz), sondern dem Ton C (65 Hz) etc. Unter dieser Voraussetzung würden sich WILLIS' Beobachtungen bemerkenswert mit den in Kapitel 6 vorgestellten eigenen Ergebnissen decken.

<sup>21</sup>vgl. hierzu die Anmerkung zu LEONHARD EULERS Vokalterminologie auf S. 34.

Mundes, der ihn zur Aussprache der Wörter geschickt macht, ist ohne Zweifel auch dieß, daß es der Geschicklichkeit des Menschen bisher noch nicht gelingen wollen, ihn durch Maschinen nachzuahmen. [...]

Ohne Zweifel wäre das eine von den wichtigsten Entdeckungen, wenn man eine Maschine erfände, die alle Töne unserer Wörter mit allen ihren Artikulationen aussprechen könnte. Wenn man jemals mit einer solchen Maschine zustande käme, und sie durch gewisse Orgel- oder Klavier-Tasten alle Wörter könnte aussprechen lassen; so würde alle Welt mit Recht erstaunt seyn, eine Maschine ganze Reden hersagen zu hören, die man mit der größten Anmuth würde vergesellschaften können. Die Prediger und Redner, deren Stimme nicht stark oder angenehm genug wäre, könnten alsdann [...] auf einer solchen Maschine spielen, so wie jetzt die Organisten musikalische Stücke spielen. Die Sache scheint mir nicht unmöglich zu seyn. « [Euler 1773, 235 ff.].

In der Mitte des 18. Jahrhunderts existierte also die Annahme, dass es »eigentlich« nicht unmöglich sein dürfe, mit mechanischen Mitteln nach Art eines technisch komplexen Musikinstrumentes wie eben der Orgel eine umfassende Sprachsynthese herzustellen. Hierbei unterlag man aber, neben der Missachtung fundamentaler Sprachphänomene wie der Koartikulation, in Bezug auf die zentrale Frage der Anregung allem Anschein nach einem Irrtum aus dem Orgelbau, der mit der dort verbreiteten Pfeifenbauweise der VH zusammenhing. Dass die assoziative Verbindung zwischen den beiden Themen Sprachsynthese und Orgelbau offenbar weit verbreitet war, zeigt auch ein Vergleich, den KRATZENSTEIN im Zusammenhang mit Sprachherzeugung bzw. deren Erlernen bemüht:

»Dès qu'un enfant, qui a des organes bien conformés, entend ses semblables articuler des mots, il cherche à les imiter; il fait mouvoir successivement les différens muscles du thorax, du larynx & de la bouche, semblable à un Organiste qui cherche sur son clavier le ton qu'il veut toucher [...].« [Kratzenstein 1782, 361]

»Wenn ein Kind, das über wohlgeformte (Sprach-)Organe verfügt, in der ersten Phase des Sprechenlernens versucht, diejenigen Wörter, die es Erwachsene artikulieren hört, zu imitieren, setzt es die Muskeln von Thorax, Larynx und Mund in Bewegung, ganz ähnlich einem Organisten, der auf seiner Klaviatur denjenigen Ton sucht, den er spielen will.«

Offenbar hielt sich die Vorstellung eines Zusammenhanges zwischen Orgelbau bzw. der VH und einer seriösen und umfassenden Sprachsynthese bis ins 20. Jahrhundert hinein. Diesen Eindruck erwecken jedenfalls W. NIEMANNs Ausführungen zur Sprachsyntheseforschung des 17. Jahrhunderts:

»Die Vervollkommnung der Orgel, besonders die Vermehrung der Stimmen [die Entwicklung neuer Pfeifenbauformen, also auch Klangfarben], wies nun den Weg, den man bei der Konstruktion von Sprechapparaten einzuschlagen hatte, zeigte doch die vox humana, daß es möglich war, dem Klang der menschlichen Stimme nahe zu kommen. Freilich, von dem einfachen Laut bis zur artikulierten Sprache war noch ein weiterer Weg. Wichtig aber war es, daß man von nun an auch in wissenschaftlichen Kreisen dem Problem Beachtung schenkte.«  
[Niemann 1920, 18 f.]

Angesichts dieser vielfältigen Bezugnahmen auf die Orgel im Allgemeinen und die VH im Speziellen verwundert es nicht, wenn eine 1779 von der Russischen Akademie der Wissenschaften formulierte Preisfrage den Zusammenhang zwischen Sprachsynthese und Orgelbau explizit betont und dessen Konkretisierung sogar einfordert (vgl. hierzu auch das Kapitel 3.2).

Wie aus dem vorher Gesagten unschwer ersichtlich ist, darf in diesem Zitat die Wendung »Vox humana« nicht mit »menschliche Stimme« ins Deutsche übersetzt werden, wie es gleichwohl mitunter geschah (z. B. Kohler 2000), sondern sie bezeichnet als *Terminus technicus* die besprochene spezielle Pfeifenbauform im Orgelbau.

Der Glaube, hinter dem als VH bezeichneten Orgelregister verberge sich der Schlüssel zu einer erfolgreichen Sprachsynthese, war ganz offenbar stark und strahlte mit den Arbeiten CHARLES WHEATSTONES, JOSEPH FABERS und ALEXANDER GRAHAM BELLS bis weit ins 19. Jahrhundert hinein (vgl. Kap. 4.3.2, 4.3.3 und 4.3.4). Alle drei griffen das grundsätzliche Konstruktionsprinzip der Sprachsynthese WOLFGANG VON KEMPELENS auf und suchten die grundlegenden Probleme, an denen KEMPELEN seinerzeit gescheitert war, mit aus ihren Forschungen resultierenden Mitteln zu lösen. Am weitesten scheint hierbei FABER 1835 mit seiner »Euphonia« gekommen zu sein, die offenbar sowohl über eine bewegliche Zunge als auch über einen formveränderlichen Rachenraum verfügte (siehe hierzu auch das Kapitel 4.3.3). Jedoch ist diese Maschine ebenso wenig erhalten wie diejenige BELLS; auch die Konstruktionen KEMPELENS [Kempelen 1791a], KRATZENSTEINS [Kratzenstein 1781] und MICALS müssen als verschollen gelten. Lediglich die KEMPELEN-Replik CHARLES WHEATSTONES existiert noch (vgl. Kapitel 3.4.2 und 4.3.2).

Auffällig ist die Beziehung, die zwischen offenbar allen seriösen Sprachsynthesekonzepten des 18. Jahrhunderts und der VH besteht. Über die Gründe hierfür kann bislang nur spekuliert werden. Dieses Phänomen pauschal mit der Literaturlage zu begründen, die einen Zusammenhang häufig nahelegte, erscheint zu kurz gegriffen. Insbesondere Personen wie KEMPELEN, WILLIS oder WHEATSTONE, die sich nachweislich kritisch-distanziert und mit umfassendem anatomischem Wissen der Sprachsynthese näherten, dürfte man wohl nicht unterstellen, derartig oberflächlichen Kolportagen tieferen Wahrheitsgehalt beigemessen zu haben.

Das Thema der VH ist im Orgelbau bis heute nicht abgeschlossen. Auch, weil das dieser Namensgebung zugrundeliegende Missverständnis selbst in Orgelbau-Fachkreisen oft unbekannt ist. Aus heutiger Sicht scheint es dennoch geradezu absurd, ausgerechnet mit solchen Orgelpfeifen eine praxistaugliche Sprachsynthese zuwege bringen zu wollen. Trotzdem »erlangen« selbst so umfassend gebildete Wissenschaftler wie KEMPELEN diesem Mythos, auch wenn sie glaubten, ihn umgangen zu haben (vgl. [Kempelen 1791, 390 ff.]).

## 2.4 Zusammenfassung

*Register* einer Orgel erhalten ihren Namen aus unterschiedlichen Gründen. Teilweise sind es Funktionsbezeichnungen wie »Oktave«, »Quinte« etc. Es können auch Benennungen sein, die die Bauform der jeweiligen Pfeifen charakterisieren wie beispielsweise »Gedeckt«, »Bourdon« etc. Eine dritte und außerordentlich vielfältige Gruppe bilden schließlich die *Register*, deren Namen eine Klangbeschreibung darstellen. Hierzu gehören beispielsweise »Flöte« oder »Gambe«. Das hier interessierende *Register* VH wurde traditionell als dieser letzten Gruppe zugehörig interpretiert, wenngleich es starke Hinweise darauf gibt, dass mit »Vox humana« ursprünglich weniger eine Charakterisierung des Klanges als des Verwendungszweckes gemeint gewesen sein dürfte.

Die Pfeifen der als VH bezeichneten *Register* gehören stets der Gruppe der *aufschlagenden Lingualpfeifen* an. Kennzeichnend für sie ist ein verhältnismäßig kurzer *Resonator*, der in keinem proportionalen Zusammenhang zum Grundton der jeweiligen Pfeife steht, wie es bei anderen *Lingualpfeifen* der Fall ist. In ihrer ursprünglichen französischen Tradition wurde die VH von dem *Lingualregister Cromorne* abgeleitet, von dem sie sich durch eben jene unproportionalen *Resonatoren* unterscheidet. Auf dem Gebiet des deutschen Orgelbaues weist die VH – anders als in ihrem Ursprungsland Frankreich – eine große Vielfalt an Bauformen auf, die ins-

besondere durch oftmals aufwändig gestaltete *Resonatoren* gekennzeichnet ist. Bei just den *Registern*, über die in einigen der im Vorangehenden angeführten Quellen hinsichtlich ihrer »Vokalähnlichkeit« positiv berichtet wurde, ist ihre konkrete Bauform heute nicht mehr bekannt. Ebenfalls muss in Ermangelung aussagekräftiger Quellen offen bleiben, in wie weit die grundlegende Veränderung der Singästhetik im 18. Jahrhundert zugunsten des *Bel canto* in Frankreich oder Deutschland frühere Singformen ablöste. Wenn es diese gegeben hat und sie der heute noch auf Sardinien praktizierten Technik des *Canto a tenore* ähnelten, würde dies die Charakterisierung der VH zumindest in Teilen erklären können.

Andere in diesem Kapitel angeführten Quellen belegen, dass man sich im 18. Jahrhundert der enormen Diskrepanz zwischen dem faktischen Klang der VH-Pfeifen und ihrem programmatisch aufgefassten Namen durchaus bewusst war. Anstatt es aber dabei bewenden zu lassen oder den Pfeifen dieser Bauform einen neuen Namen zu verleihen, versuchten zumindest einzelne Orgelbauer – und in ihrer Folge auch Sprachsynthesepioniere – die Pfeifen so weiterzuentwickeln, dass sie ihrem Namen besser gerecht wurden. Durch das gesamte 18. und auch 19. Jahrhundert zieht zudem die rational-naturwissenschaftliche Beschäftigung mit dem »Mysterium« der VH. Dies zeigen die angeführten Zitate von DODART, EULER und WILLIS.

Im Folgenden sollen die zahlreichen Sprachsynthesekonzepte, die im 18. Jahrhundert entstanden, geschildert werden, wobei das Phänomen der VH geradezu einen roten Faden darstellt.

# Kapitel 3

## 18. Jahrhundert: Grundlegende Konzepte

Angeichts der Vielzahl von seriösen Ansätzen zur Sprachsynthese, die in den Jahren zwischen 1770 bis 1791 geradezu aus dem Boden schoss (mindestens vier verschiedene und vollkommen autonome Ansätze in nur etwa 20 Jahren!), scheint die Zeit regelrecht reif gewesen zu sein für derartige Versuche. Alle vier Urheber lebten und forschten in vollkommen unterschiedlichen Gegenden Europas und es scheint, als ob sie jeweils beinahe nichts voneinander und ihrem Bemühen gewusst hätten. Nur für WOLFGANG VON KEMPELEN ist durch sein eigenes Zeugnis belegt, dass er die Arbeiten KRATZENSTEINS zumindest zur Kenntnis genommen hatte [Kempelen 1791a, 197 ff.]. ERASMUS DARWIN hatte gerüchtehalber gehört, dass außer ihm noch andere an der Lösung des Problems arbeiteten. Er schrieb 1772 an BENJAMIN FRANKLIN:

»I have heard of somebody that attempt to make a speaking machine, pray was there any Truth in such Reports?«  
[Darwin 1981, 63]

### 3.1 Erasmus Darwin

Die Überlieferung zur Sprachmaschine ERASMUS DARWINS (1731–1802) ist noch dürttiger als bei seinen Zeitgenossen MICAL oder KRATZENSTEIN. Lediglich DARWIN selbst beschreibt in seinem 1803 posthum erschienenen Grundlagenwerk »*The Temple of Nature*« den Aufbau der Maschine:

»...I contrived a wooden mouth with lips of soft leather, and with a vale back part of it for nostrils, both which could be quickly opened or closed by the pressure of the fingers, the vocality was given by a silk ribbon about an inch long and a quarter of an inch wide stretched between two bits of smooth wood a little hollowed; so that when a gentle current of air from bellows was blown on the edge of the ribbon, it gave an agreeable tone, as it vibrated between the wooden sides, much like a human voice. This head pronounced the p, b, m, and the vowel a, with so great nicety as to deceive all who heard it unseen, when it pronounced the words mama, papa, map, pam; and had a most plaintive tone, when the lips were gradually closed. My other occupations prevented me from proceeding in the further construction of this machine; which might have required but 13 movements, as shown in the above analysis, unless some variety of musical note was to be added to the vocality produced in the larynx; all of which movements might communicate with the keys of a harpsichord or forte piano, and perform the song as well as the accompaniment; or which if built in a gigantic form, might speak so loud as to command an army or instruct a crowd.« [Darwin 1803, 119 f.]

Erste private Vorführungen dieser Apparatur gab es offenbar bereits 1770, ein Jahr darauf dann öffentliche in der Lunar Society in Birmingham [Jackson 2005, 229]. Diese scheinen erheblichen Eindruck gemacht zu haben, denn im selben Jahr schloss MATTHEW BOULTON (1728-1809), Partner von JAMES WATT und prägendes Mitglied der Lunar Society, einen Vertrag mit DARWIN:

»I promise to pay to Dr. Darwin, of Lichfield, one thousand pounds upon his delivering to me (within two years from the date hereof) an instrument called an organ that is capable of pronouncing the Lord's prayer, the Creed and Ten Commandments in the vulgar tongue, and his ceding to me, and me only, the property of the said invention with all the advantage thereto appertaining.« [Jackson 2005, 231]

Zu dieser Anfertigung scheint es jedoch nie gekommen zu sein. Bemerkenswert ist die Bezeichnung der Maschine als »Orgel« und die Festlegung eines genauen Repertoires. Möglicherweise war eine Steuerung über *Stiftwalzen* wie bei MICAL vorgesehen (vgl. Kap. 3.3).



Zeichnungen oder präzise Beschreibungen und Aussagen über die Synthesequalität von dritter Hand existieren offenbar nicht; lediglich ein retrospektiver Ohrenzeugenbericht aus dem Jahre 1798 ist überliefert:

»The speaking machine, which is just announced from France<sup>22</sup> does not say so many words as yours did many years ago. It prattles only papa and mama. Yours spoke those words, and could also say go. I placed one of your mouths in a room near some people in 1770, who actually thought I had a child with me, calling papa and mama.« [Edgeworth 1820, 198 f.]

Anhand dieser Beschreibungen kann davon ausgegangen werden, dass es sich bei DARWINS sprechendem Kopf um einen ziemlich direkten Vorläufer der Arbeiten MICALS und vor allem KEMPELENS gehandelt haben muss. Auffällige Parallelen zu letzterem bestehen in der Konstruktion der »Nase« und insbesondere im Vokabular der Vorrichtung, auch wenn »Papa« und »Mama« als Äußerungen grundsätzlich naheliegen, da CV-Silben einfach zu realisieren sind. Die Idee der Steuerung über Tasten findet sich späterhin sowohl bei KEMPELEN als auch bei EULER, KRATZENSTEIN und auch MICAL.

## 3.2 Christian Gottlieb Kratzenstein

### 3.2.1 Biografisches

Der Naturforscher KRATZENSTEIN (1723–1795) verkörpert einen außergewöhnlichen Fall von fachübergreifender wissenschaftlicher Kompetenz. Seine nachmalig populären Forschungen zur Sprachsynthese führten zwar auf diesem Gebiet in die Irre und damit auch in eine technologische Sackgasse, jedoch legten sie zugleich im Bereich des Orgelbaues den Grundstein für eine das gesamte 19. Jahrhundert andauernde Erfolgsgeschichte. Der folgende biografische Abriss folgt insbesondere der Darstellung bei SPLINTER [2005].

Ähnlich wie im Falle WOLFGANG VON KEMPELENS kursiert auch über KRATZENSTEIN eine Vielzahl von Fehlzuschreibungen und Legenden. Geboren wurde er am 30. Januar 1723 als Sohn eines Oberschullehrers (oder auch Bürgermeisters oder Juristen) in Wernigerode. Seinem offenbar schon in Kindertagen regen Interesse an naturwissenschaftlichen Fragen folgend,

<sup>22</sup>BUSCH zufolge handelte es sich um eine ca. 70 cm große androidenartige Puppe mit einem Trichter vor dem Mund, in den Fragen hineingesprochen werden konnten (!), worauf diese antwortete [Busch 1817, 496].

studierte er ab 1742 an der Universität Halle (Saale) Naturwissenschaft und promovierte dort 1746 zum Dr. med., um dortselbst direkt anschließend als Privatdozent Medizin und Naturlehre zu unterrichten. Bereits 1748 folgte er auf Vermittlung von LEONHARD EULER (1707–1783) einem Ruf an die Russische Akademie der Wissenschaften in St. Petersburg und war seit demselben Jahr auch Mitglied der Deutschen Akademie der Naturforscher (Leopoldina). Nach fünf Jahren als Professor für Mathematik und Mechanik in St. Petersburg wechselte er schließlich 1753 auf eine Professur für Experimentalphysik nach Kopenhagen, die er nunmehr 33 Jahre lang bekleiden sollte. Nicht zuletzt durch seine umfangreiche Korrespondenz mit Wissenschaftskollegen besaß KRATZENSTEIN einen weit verbreiteten und sehr guten Ruf als Naturforscher im Allgemeinen und Mechaniker im Besonderen.

Seine Interessen waren – nicht ungewöhnlich für einen naturwissenschaftlichen Gelehrten seiner Zeit – außerordentlich breit gefächert. Er befasste sich mit der Möglichkeit der Luftschifffahrt, mit Aspekten der Alchemie, Astronomie, Metereologie, Navigation und insbesondere mit der Wirkung von Elektrizität auf den menschlichen Körper. Hierfür entwickelte er eine der ersten Elektrisiermaschinen und wandte auch bereits eine Form der Elektrotherapie beim Menschen an. Sein Beitrag zur Sprachsyntheseforschung mutet vor diesem Hintergrund als eine eher nachrangige Episode an.

Einen Monat vor seinem Tod wurde 1795 KRATZENSTEINS Kopenhagener Haus Opfer des verheerenden Stadtbrandes, bei dem er neben beinahe aller anderen Habe auch seine bedeutende Sammlung alter Handschriften und physikalischer Instrumente sowie seine gesamte Korrespondenz verlor.

### 3.2.2 Leonhard Euler & Kratzensteins »Tentamen«

Bereits ab 1747, also ein Jahr vor seinem Ruf nach St. Petersburg, stand KRATZENSTEIN in regem brieflichen Kontakt mit LEONHARD EULER. Auch mit anderen namhaften Gelehrten wie JOHANN II. BERNOULLI (1710–1790) oder CARL VON LINNÉ (1707–1778) korrespondierte er lebenslang. Der durch den Austausch mit EULER entstandene, stark vom Orgelbau beeinflusste Prototyp einer Vokalsynthese sollte zwar nicht für die Sprachsyntheseforschung, wohl aber für den Orgelbau von großer Bedeutung werden.

Der heute u. a. durch die *Eulersche Zahl* bekannte Mathematiker LEONHARD EULER betrieb allem Anschein nach auch akustische Forschungen, die sowohl im Bereich der Musiktheorie als auch in der (freilich damals so noch nicht existenten) artikulatorischen Phonetik deutliche Spuren hinterlassen sollten. Bereits 1773 hatte er in seinen »*Briefen an eine deutsche*

*Prinzessinn über verschiedene Gegenstände aus der Physik und Philosophie*«<sup>23</sup> zunächst seine Vorstellungen vom Schall im Allgemeinen und der menschlichen Stimme im Besonderen entwickelt, um anschließend die Idee einer Sprachmaschine zu präsentieren, die ähnlich dem Funktionsprinzip einer Orgel konstruiert sein sollte.

Bemerkenswert sind seine Beobachtungen zur Frage, wie es zu unterschiedlichen Vokalqualitäten komme, und der daraus folgende Schluss, der zu dieser Zeit noch bei weitem kein Allgemeingut darstellte:<sup>24</sup>

»Die wunderbarste Verschiedenheit bemerkt man an der menschlichen Stimme, die das größte Meisterstück des Schöpfers ist; ohne noch von den verschiedenen Artikulationen zu reden, aus welchen die Rede besteht. Geruhen Ew. H. nur über die verschiednen Vocalen nachzudenken, die der Mund ganz einfach ausspricht oder singet. Wenn man den Buchstaben a singt oder ausspricht, so ist der Ton ganz anders, als wenn man den Buchstaben e oder o oder i oder u oder ai u. s. w. ausspräche oder sänge, auch wenn man bey allen in einerley Tone bleibe. Man kann also die Ursache dieses Unterschiedes nicht in der Schnelligkeit oder der Ordnung der Schwingungen suchen; sie ist uns verborgen, und die Philosophen haben sie bisher noch nicht ergründen können.

Ew. H. werden leicht gewahr werden, daß man zur Aussprache dieser verschiednen Vocalen der Höhlung des Mundes eine verschiedne Gestalt geben muß, und unser Mund ist dazu vor dem Munde aller Thiere vorzüglich eingerichtet.« [Euler 1773, 235]

Zwei Dinge sind hier bemerkenswert: Zum einen EULERS Ansicht, es sei Aufgabe der Philosophen, die physiologische Grundlage der unterschiedlichen Vokalqualitäten zu ergründen. Dies erklärt sich freilich dadurch, dass die Philosophie seinerzeit noch grundsätzlicher als heute als »Mutter aller Wissenschaften« galt und somit auch essentieller, ja maßgeblicher Bestandteil der Disziplinen Medizin und Anatomie war. Zum anderen haderte auch noch WOLFGANG VON KEMPELEN mit dem subjektiven Eindruck, dass

<sup>23</sup>Diese sind entgegen der häufig zu lesenden Auffassung keine fiktiven Briefe (wie sie in der Gattung der *offenen Briefe* durchaus häufig vorkommen), sondern waren an Friederike Charlotte von Brandenburg-Schwedt (1745–1808), nachmalig die letzte Fürstäbtissin des freiweltlichen Damenstifts Herford, gerichtet [Anonymus 1877, 424].

<sup>24</sup>EULER zeigt sich hierbei auch als Gegner der Theorien GOTTFRIED HERDERS (1744–1803), da er von einem göttlichen Ursprung der Sprache ausgeht. Um so bemerkenswerter ist sein Wunsch, diese göttliche Gabe mit menschlicher Ingenieurskunst nachzubilden.

sich Vokale hauptsächlich durch ihre Tonhöhen unterschieden und löste sich erst spät und widerstrebend von dieser Vorstellung [Kempelen 1791a, 196 und 400]. Die Idee, auf Grundlage einer Orgel als mechanischem Steuerungsgerät und der VH als Stimmanalogon eine Sprachsynthese zu entwickeln, legt EULER im direkten Anschluss an das obenstehende Zitat dar (vgl. Zitat auf S. 24) [Euler 1773, 235 ff.].

Zunächst irritierend ist die Verwendung der Diphthonge suggerierenden Buchstabenkombinationen *ae* und *ai* in beiden EULER-ZITATEN. Es erscheint aber als ausgeschlossen, dass hier von Diphthongen im heutigen Sinne die Rede ist. Eine analoge Stelle aus EULERS posthumer Schrift »*Meditatio de formatione vocum*« bietet hier einen Anhaltspunkt: Hier wird die lateinische Buchstabenkombination »*ae*« mit der griechischen Buchstabenkombination »*α*« gleichgesetzt [Euler 1862, 798]. Diese wird als [ɛ] ausgesprochen, so dass diese Lautqualität wohl auch für das »*ae*« in den »*Briefen*« anzunehmen ist. Hinsichtlich des »*ai*« zeigt ein Vergleich der deutschen mit der (originalen) französischen Fassung, dass dem Übersetzer oder Drucker hier ein Fehler unterlief: In der französischen Ausgabe ist an dieser Stelle von den Vokalen »*ai ou œ*« (also [ɛ] und [œ]) die Rede [Euler 1768, 97], womit die in Kapitel 6.3 vorgestellten eigenen Forschungsergebnisse gut harmonieren. Auch KRATZENSTEIN umschreibt später die von seiner Konstruktion hervorgebrachten Laute auf diese Weise und bezeichnet sie sogar explizit als Diphthonge [Kratzenstein 1782, 379 f.]. Gemeint sein dürften allerdings auch hier Umlaute.

Bemerkenswert ist hier auch EULERS Schlussfolgerung aus der Tatsache der enormen Komplexität der Sprachartikulation, die von zahlreichen Zeitgenossen wie dem Theologen JOHANN PETER SÜSSMILCH (1707-1767), dem Philosophen BERNARD LAMY (vgl. Kapitel 1) u. a. gerade als Argument für den göttlichen Ursprung der Sprache ins Feld geführt wurde. In dem kurzen, erst posthum veröffentlichten Aufsatz »*Meditatio de formatione vocum*« [Euler 1862, 798 f.] legt EULER sehr prägnant seine Theorie des Zusammenhanges zwischen der VH und einer erfolgreichen Sprachsynthese dar: Den bisherigen VH in Orgeln attestiert er, nicht die Sprache an sich nachzubilden, sondern nur den Effekt, den eine singende Stimme hervorbringe, was ein Leichtes sei. Die Synthese von Konsonanten sei im Vergleich zu derjenigen von Vokalen ungleich schwieriger zu bewerkstelligen. Schon jetzt aber sei zu beobachten, dass bei manchen VH-Registern höhere Töne eher nach einem Ä-ähnlichen Laut klängen, tiefere O-ähnlich oder wie ein kurzes (ungespanntes) U.<sup>25</sup> Offenbar hänge viel von der Gestaltung des *Resonators* ab, wenn es gelte, einen definierten Vokal zu erzeugen. Wenn

---

<sup>25</sup>Vgl. hierzu Kap. 6.3.

es gelänge, die Vorgänge im Mund zu verstehen, die zur Artikulation der verschiedenen Vokale führten, so sei es ein Einfaches, dies auf die Resonatoren der VH zu übertragen.<sup>26</sup> Der grundsätzliche Gedanke, dass die Orgel zum Vorbild einer Sprachsynthese dienen könne, war indessen nicht neu (vgl. Zitate in Kap. 2.3).<sup>27</sup>

Im Jahre 1779 stellte nun die Petersburger Akademie eine Preisfrage, deren auf EULERS Überlegungen maßgeschneiderte Formulierung auffällt:

»1) Qualis sit natura et character sonorum litterarum vocalium a, e, i, o, u tam insigniter inter se diversorum. 2) Annon construi queant instrumenta ordini tuborum organicorum, sub termino vocis humanae noto, simila, quae litterarum vocalium a, e, i, o, u sonos exprimant.« [Kratzenstein 1781, Deckblatt]

»1) Was sind die Natur und der Charakter der Vokal-Buchstaben a, e, i, o, u die sich so deutlich voneinander unterscheiden? 2) Ist es nicht möglich, Instrumente nach Art derjenigen Orgelpfeifen zu bauen, die als Vox humana bekannt sind und die den Klang der Vokal-Buchstaben a, e, i, o, u hervorbringen?«

Die Idee zu dieser Preisfrage stammt in der Tat von EULER und ist sogar noch weitaus älter als die entsprechende Thematisierung in seinen »*Briefen*«: Am 12. Mai 1765 übersandte er auf Bitten des Professors für Eloquenz (!) JACOB VON STAEHLIN (Jakow Jakowlewitsch Staehlin, 1709–1785) diesem einen Katalog von 23 Fragen zu von ihm als virulent empfundenen und noch nicht befriedigend gelösten Problemen in den Naturwissenschaften als Vorlage für künftige Preisfragen der Petersburger Akademie. Bereits EULERS zweite Frage dort lautet:

»Cum Theoria soni jam ita sit exulta, ut ratio sonorum gravium et auctorum, que fundamentum harmoniae continetur, satis perspicue intelligatur, aliaque sonorum differentiae

<sup>26</sup>Exakt diese Argumentationskette verfolgte auch KRATZENSTEIN in seinem »*Tentamen*« (vgl. Kap. 7.1).

<sup>27</sup>Noch im 20. Jahrhundert wurde die Idee vom Komponisten GYÖRGY LIGETI (1923–2006) propagiert: »Meine Lieblingsidee ist ein Sprachwerk, die sprechende Orgel. Beim Sprechen übernehmen die Stimmbänder, die eigentlich Doppelzungen [wohl: Doppelrohrblatt] sind, die Rolle des Tongenerators. Durch Resonanzänderung der Mund- und Nasenhöhle usw. können verschiedene Teiltöne hervorgehoben bzw. unterdrückt werden, und es entstehen die durch ihre typischen Formanten unterschiedenen Vokale. Außer den Stimmbändern als Tongeneratoren besitzen wir Geräuschgeneratoren zwischen Lippen, Zunge und Zähnen oder Gaumen, Zäpfchen usw., die die Konsonanten erzeugen. Das alles könnte man mechanisch nachbauen.« [Eidenbenz 2006, 43]

Auch der Arzt FRITZ KAHN (1888–1968) stellte auf einem seiner populärwissenschaftlichen Sagittalschnitte die Glottis als Orgel dar, die für jeden Buchstaben (!) eine einzelne Pfeife besitzt [Debschitz & Debschitz 2013, 185].

explicari possint; ea sonorum diversitas, qua vocales a, e, i, o, u efferentur, adhuc prosus est incognita, unde indoles sonorum, quibus singulae vocales enunciantur, explicanda proponitur, simulque quaeritur, annon ejusmodi instrumenta confiri possint, quae sonos vocalium perfecte imitentur.« [Juškevič & Winter 1976, 236]

»Obgleich die Lauttheorie bereits so weit ausgebildet ist, dass das Prinzip der tiefen und langen Laute, welche das Fundament der Harmonie beinhalten, ausreichend deutlich verstanden wurde, um die Unterschiede zwischen diesen Lauten erklären zu können, ist diese Verschiedenheit der Laute, die von den Vokalen a, e, i, o, u hervorgebracht wird, bisher geradezu unbekannt, weshalb die natürliche Beschaffenheit der Vokale zu erklären vorgeschlagen wird und zugleich zu fragen ist, ob Instrumente solcher Art hergestellt werden können, dass diese Vokallaute gänzlich nachgeahmt werden.«

Zwei weitere Fragen aus EULERS Katalog bildeten in bemerkenswerter Regelmäßigkeit die Preisfragen der Petersburger Akademie für die Jahre 1777 (Nr. 4) und 1783 (Nr. 11) [Juškevič & Winter 1976, 238, Fußnote 2]. Der Hinweis auf die Orgel als »Syntheseprototyp« oder gar die VH fehlt in dieser Urfassung der Fragestellung. Möglicherweise hat sich dieser Zusammenhang bei EULER erst im Laufe der nächsten Jahre herausgebildet, so dass sich in den »*Briefen*« bereits der deutliche Verweis auf die Orgel, bei der endgültigen Fassung der Preisfrage dann auch derjenige auf die VH findet.

Da die Korrespondenz zwischen EULER und KRATZENSTEIN auf Seiten KRATZENSTEINS verloren und aufseiten EULERS nur bis in die 1750er Jahre überliefert ist, kann eine Absprache hinsichtlich der Preisfrage nicht belegt werden. Zudem war EULER nach 1771 vollständig erblindet. Seine Korrespondenz erledigte statt seiner sein ältester Sohn JOHANN ALBRECHT EULER (1734–1800), der zugleich ständiger Sekretär an der Petersburger Akademie war. In deren Archiv sind vier Briefe KRATZENSTEINS an J. A. EULER aus den Jahren 1780/81 erhalten, in denen die Beförderung der »Vokalmaschine« und die Veröffentlichung des »*Tentamen*« thematisiert wird. Zusammenfassungen dieser auf Deutsch geschriebenen Korrespondenz sind in russischer Sprache publiziert [Ljubimenko 1937, 399, 346 f. und 351].<sup>28</sup> Leider war es trotz intensiver Bemühungen nicht möglich, Kopien oder Abschriften der originalen Brieftexte zu erhalten.<sup>29</sup>

<sup>28</sup>Für die Übersetzung dieser Texte habe ich Frau Dr. ANNA VOLODINA herzlich zu danken.

<sup>29</sup>Hieran vermochte sogar der persönliche Besuch von Frau Prof. MARIJA SMIRNOVA (Moskau), im Archiv der Russischen Akademie der Wissenschaften nichts zu ändern.

Bemerkenswerterweise enthält die französische Fassung der Preisfrage im Brief KRATZENSTEINS an JOHANN ALBRECHT EULER vom 9. Januar 1781 keinen expliziten Hinweis auf die VH und ähnelt damit stark der Urfassung der Fragestellung:

»Quels sont le nature et la caractère des tons vocaux, si essentiellement différens entre eux, et ne pourrait-on pas faire des instruments, semblables aux tyaux d'orgues, qui les imitent parfaitement?« [Ljubimenco 1937, 351]<sup>30</sup>

»Was sind die Natur und der Charakter der Vokaltöne, die sich so grundlegend voneinander unterscheiden und wäre es nicht möglich, Instrumente wie aus Orgelpfeifen zu bauen, die diese perfekt nachahmen?«

Es gibt keinerlei Hinweise darauf, dass einer von KRATZENSTEINS »Konkurrenten« DARWIN, MICAL und KEMPELEN rechtzeitig von dieser Preisfrage erfahren hätte. Für die These, dass der stets von ihm wertgeschätzte Kollegen fördernde EULER nicht nur hinsichtlich der Fragestellung, sondern bezüglich des künftigen Preisträgers großen Einfluss hatte, spricht ein ähnlicher Fall aus dem Jahre 1746: Damals hatte die Berliner königlich-preußische Akademie der Wissenschaften eine Preisfrage zum Wesen des Windes gestellt. Obwohl die Antworten eigentlich anonym eingereicht und bewertet werden sollten, forderte EULER hier zunächst DANIEL BERNOULLI (1700–1782) persönlich zu einer Einreichung auf, später dann JEAN-BAPTISTE LE ROND D'ALEMBERT (1717–1783), der schließlich den Preis auch erhielt [Kleinert 1989]. Im Jahre 1749 war EULER der Spiritus rector hinter der Preisfrage der Petersburger Akademie. Auch wenn er explizit ersucht wurde, seine Urheberchaft der Preisfrage nicht offenzulegen, war es scheinbar trotzdem ein offenes Geheimnis, dass EULER bei diesem Wettbewerb nicht nur als Juror, sondern auch als Autor fungierte [Juškevič & Winter 1961, 10].

Nach WILLIS hatte sich KRATZENSTEIN bereits seit etwa 1770, also zeitgleich mit KEMPELEN, mit Fragen zur Sprachsynthese befasst [Willis 1832, 398].<sup>31</sup> EULER hatte ihn (wie auch andere Nachwuchswissenschaftler, die seinen strengen Kriterien entsprachen) bereits seit den 1750er Jahren nachhaltig protegiert und besaß eine hohe Meinung von seinen Fähigkeiten als Mechaniker und »Instrumentenmacher«.<sup>32</sup> All diese Indizien deuten stark

<sup>30</sup>Die publizierte frz. Fassung von KRATZENSTEINS Aufsatz erwähnt indessen die VH wieder explizit in der Preisfrage [Kratzenstein 1782, 359].

<sup>31</sup>Willis bezieht sich hierbei vermutlich auf eine Stelle im »*Tentamen*«, die diese Annahme nahelegt [vgl. Kratzenstein 1781, 40].

<sup>32</sup>EULER am 25.06/06.07.1748 an JOHANN DANIEL SCHUMACHER (1690–1761): »Zu der Mechanic hätte in der That keine tüchtigere Person gefunden werden können, indem er [Kratzenstein] nicht nur eine vollkommene Erkenntniß von allen Maschi-

# CHRISTIANI THEOPHILI KRATZENSTEINII

MEICINAE DOCTORIS, PROFESSORIS PHYSICES PU-  
BLICI ORDINARI IN UNIVERSITATE HAFNIENSI,  
ACADEMIAE SCIENTIARUM IMPERIALIS PE-  
TROPOLITANAE MEMBRI,

## T E N T A M E N

RESOLVENDI PROBLEMA AB ACADEMIA SCIEN-  
TIARUM IMPERIALI PETROPOLITANA AD AN-  
NUM 1780 PUBLICE PROPOSITUM.

- 1) *Qualis sit natura et character sonorum litterarum vocalium  
a, e, i, o, u tam insigniter inter se diversorum.*
- 2) *Annon constitui queant instrumenta ordini tuborum organi-  
corum, sub termino vocis humanae noio, similia, quae  
litterarum vocalium a, e, i, o, u, sonos exprimant.*

---

In  
Publico Academiae Conventu  
die XIX. Septembris MDCCCLXXX.

Praemio coronatum

---

*Plus ultra.*

---

P E T R O P O L I  
T Y P I S A C A D E M I A E S C I E N T I A R V M  
M D C C L X X X I.

Abbildung 3.1: Titelblatt von KRATZENSTEINS »Tentamen« (1781).

auf eine von EULER arrangierte Ausschreibung des Wettbewerbs »ad personam« hin.

Bereits 1781, nur ein Jahr nach Veröffentlichung der Preisfrage, reichte KRATZENSTEIN seine Schrift ein, deren barocker Titel bereits beinahe einer Inhaltsangabe gleichkommt (Abb. 3.1).<sup>33</sup>

Für seine Ausführungen bekam KRATZENSTEIN den ersten Preis (»plus ultra«) zugesprochen. Im ersten Teil des »Tentamen« befasste er sich for-

---

nen besitzt, sondern auch bey einem jeden vorkommenden Fall die bequemste Art sehr sinnreich ausfündig zu machen und auch selbst ins Werk zu richten imstande ist [...].« [Juškevič & Winter 1961, 136 f.] Und am 9./20.11.1751: »An des H. Kratzensteins Schriftt habe ich mich auf gantz ungemeine Art ergötzet und seine Scharfsinnigkeit in mechanischen Erfindungen nicht genugsam bewundern können.« [Juškevič & Winter 1961, 260]

<sup>33</sup>Offenbar war KRATZENSTEINS Originaltext auf Französisch geschrieben und wurde erst für die Veröffentlichung der Petersburger Akademie ins Lateinische übertragen [vgl. Ljubimenko 1937, 351]. Ob die 1782 in den »Observations sur la physique...« erschienene französische Fassung der Abhandlung KRATZENSTEINS Originaltext oder eine weitere Bearbeitung desselben darstellt, konnte nicht ermittelt werden.



derungsgemäß mit dem Wesen und der Produktion der fünf Vokale A, E, I, O und U (»De genesi vocalium«). In diesem Teil findet sich abschließend auch eine tabellarische Synopse der artikulatorischen Prozesse, die nach KRATZENSTEIN zur Hervorbringung der fünf Vokale notwendig seien (Tab. 3.1). Indessen erhellt selbst ein Vergleich der lateinischen und der französischen Version des Tabellentextes nicht in jedem Detail seine Intentionen.

**Tabelle 3.1:** Artikulationsstellungen für die Vokale A, E, I, O, U (Übersetzung nach [Kratzenstein 1781, 15] und [Kratzenstein 1782, 364]).

	Larynx	Zunge	Gaumen- Öffnung	Öffnung d. Zähne	Lippen- Öffnung
A	Die Seiten desselben ein wenig abgesenkt und gewölbt. Der Kehldeckel ein wenig angehoben.	Die Spitze liegt an der Wurzel der Zähne des Unterkiefers. Der Zungenrücken etwas gehoben	$\frac{2}{3}$ '' <sup>34</sup>	$\frac{1}{3}$ ''	Höhe 5''' Breite 18'''.
E	Die Articulatio <sup>35</sup> des Kehldeckels wird ein wenig angehoben und zurückgezogen.	Die Spitze an die Schneide der unteren Zähne. Der Zungenrücken stärker angehoben.	$\frac{1}{3}$ ''	$\frac{1}{6}$ ''	4''' 18'''.
I	Dieselbe Position, aber die Articulatio ist flacher. Der Kehldeckel und der Rand der Glottis sind höher.	Die Zungenspitze zur Mitte der oberen und unteren Zähne oder ein wenig zwischen die Zähne.	$\frac{1}{6}$ ''	$\frac{1}{12}$ ''	2''' 18'''.

Fortsetzung auf S. 40

<sup>34</sup>Je nach zugrundeliegendem Fußmaß (Duodezimal- oder Dezimalsystem) 1''' ≈ 2,2 mm und 1'' ≈ 25 mm.

<sup>35</sup>Auch für linguistisch gebildete souveräne Kenner der jeweiligen Sprache ist weder aus der lateinischen noch aus der französischen Fassung des Textes ersichtlich, auf welche physiologische Gegebenheit Kratzenstein hier tatsächlich abhebt. Denkbar sind beispielsweise die Annahme einer Art Gelenk der Epiglottis oder aber ein Hinweis auf die (subjektive) Tonhöhe der einzelnen Vokale.

		Der Zungenrücken am stärksten angehoben. Die Oberseite formt eine Rinne, durch die die Stimme geleitet wird.			
O	Ziemlich die gleiche Position wie bei A.	Annähernd dieselbe Lage wie bei A, wobei die Zunge $\frac{1}{12}$ " mehr zurückgezogen und angehoben wird.	$\frac{1}{2}$ "	$\frac{5}{12}$ "	$3 \frac{1}{2}$ '" 8'''.
U	Die Öffnung des Kehldeckels und die Flachheit der Articulatio etwas weniger als bei I, ohne merkliche Hebung.	Die Zungenspitze ein wenig mehr als bei O von den unteren Zähnen zurückgezogen. Der Zungenrücken in seinem hinteren Teil stark angehoben.	$\frac{1}{3}$ " bis $\frac{5}{12}$ "	$\frac{1}{3}$ "	$2 \frac{1}{2}$ '" 5'''.

Als zentrales Element der Phonation glaubte KRATZENSTEIN die Epiglottis ausgemacht zu haben. Sie fungiere als eine Art Deckel des Kehlkopfes und öffne und schließe die »äußere Glottis« je nach Bedarf. Der Ansatz, dem Kehldeckel eine entscheidende Rolle bei der Stimmgebung zuzusprechen, war zu jener Zeit ebenso altbekannt wie umstritten und sollte dies auch noch über viele Jahrzehnte hin bleiben. So heißt es in »*Gehler's physikalischem Wörterbuch*« noch 1836:

»Dem Kehldeckel spricht Haller [Albrecht von Haller, 1708-1777, Schweizer Mediziner] allen Einfluss ab [...]; allein Meckel [wohl: Philipp F. Th. Meckel, 1755-1803, deutscher Mediziner und Anatom] zeigt aus anderen Versuchen, dass die Höhe und Tiefe der Töne allerdings durch das Herabgehn der Epiglottis in dem Raum über dem Kehlkopfe und durch das Anspannen ihrer Muskeln bedingt wird, auch hat Grenié [Gabriel Joseph Grenié (1756-1837), französischer Orgelbauer, verhalf der durchschlagenden Zunge zu ihrem Durchbruch im Orgelbau] durch ein ihr nachgebildetes Stückchen Papier eine Modification der Töne bei seinen Zungenpfeifen hervorgebracht. [...]

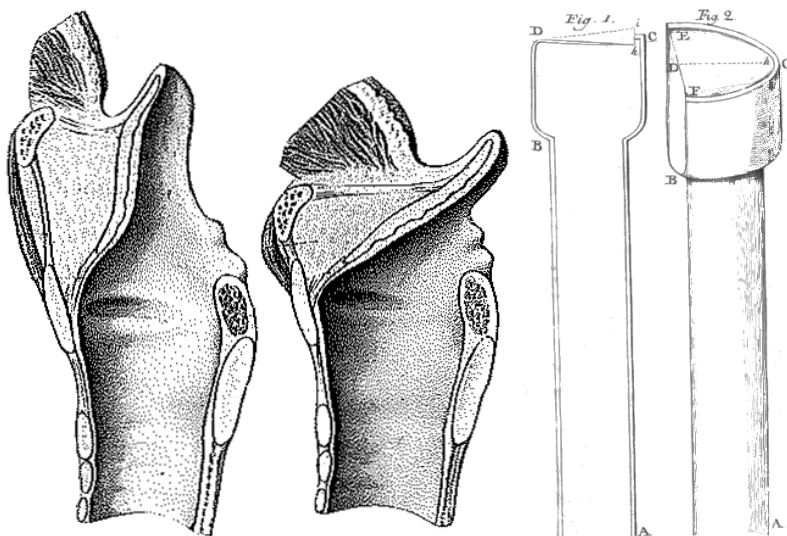
Kratzenstein ist gleichfalls der Meinung, dass die durch die Stimmritze erzeugten Töne durch den Kehledeckel, die Wölbung des Gaumens, die Zähne und auch die Lippen bedeutend modificirt werden. [...] Diese Theorie hat allerdings inneren Zusammenhang und findet in den Versuchen von Grenié, wonach solche Pfeifen, die in ihrem Baue grosse Aehnlichkeit mit dem menschlichen Sprachorgan haben, die Töne des Gesanges ziemlich genau nachbilden, eine bedeutende Unterstützung [...]. Entscheidende Gründe, die dieselbe vollständig zu widerlegen vermöchten, giebt es zwar nicht, wohl aber sind auch andere Hypothesen aufgestellt worden, die auf gleiche oder grössere Angemessenheit Ansprüche machen. Liskovius [Karl Friedrich Salomon Liscovius, 1780–1844, deutscher Arzt und Physiologe] gesteht zu, dass die Stimmbänder beim Singen zittern, allein er betrachtet dieses nicht als Ursache, sondern als Folge des Tönens [...].<sup>36</sup> [Gehler 1836, 377 f.]

Verursacht wurde diese aus heutiger Sicht abwegige Vorstellung durch eine seinerzeit populäre Theorie, die das Gewebe des Sprechapparates in durchblutetes, aktives Muskelgewebe und passives Bindegewebe unterteilte. Hinsichtlich der Stimmlippen geriet diese Theorie jedoch ins Wanken, da deren Spannung, wenn man sie als wesentlich für die Stimmgebung ansah, aktiv hätte vom Sprecher verändert werden müssen, es handelte sich aber qua definitionem um kein »aktives« Muskelgewebe. Daher nahm man an, die Stimmlippen würden passiv durch das umgebende Muskelgewebe gesteuert [Gessinger 1994, 559]. Allerdings wurden bereits 1693 die physiologischen Rollen von Glottis und Epiglottis durch SAMUEL REYHER korrekt dargestellt [Reyher 1693, 10 f.].

Schwer wurde es für KRATZENSTEIN auch, die geforderte Analogie zwischen menschlichem Stimmapparat und der VH aufzuzeigen. Er kam zu dem Schluss, dass die VH zwar im Grundsatz ein Analogon der menschlichen Anatomie sei, allerdings sei die Anordnung des Schallerzeugers eben nicht gleich; beim Menschen befindet er sich am Ende der Luftröhre, bei der VH jedoch am Ende des Ansatzrohres. Um diesen Widerspruch zu lösen, postulierte er, dass die Epiglottis als »äußere Glottis« eine andere als die bisher vermutete Funktion haben müsse und nicht nur dem Verschließen der Luftröhre dienen könne. Sie entspreche vielmehr dem Zungenblatt der Lingualpfeife. Da es in dieser aber kein Analogon zur »inneren Glot-

---

<sup>36</sup>Bemerkenswert ist hier die Verwendung des Terminus »Stimmbänder«, denen doch keine aktive Beteiligung an der Stimmproduktion zugesprochen wird.



**Abbildung 3.2:** Lage und Positon der Epiglottis beim Menschen in Atem- bzw. Schluckstellung [nach Sobotta 1959, 118] und KRATZENSTEINS Schall-generator nach dem vermeintlichen Vorbild der waagrecht in der Trachea schwingenden Epiglottis [nach Kratzenstein 1782, Tab. 1].

tis«, also zu den Stimmlippen gibt, könnten diese keinen zentralen Anteil an der Stimmgebung besitzen [Kratzenstein 1782, 368] (vgl. Abb. 3.2).

KRATZENSTEIN erkannte jedoch auch, dass die beim Menschen gegebene Abfolge von Luftröhre – Kehlkopf – Epiglottis in der VH (also in allen *Lingualpfeifen*) genau umgekehrt ist. Versuche, die Anordnung derjenigen des Menschen anzupassen, seien indes bislang stets fehlgeschlagen. Durch die neu erfundenen *durchschlagenden Lingualpfeifen* sei es nun aber möglich, dieses Problem zu lösen und die Anordnung der Teile derjenigen beim Menschen anzupassen [Kratzenstein 1782, 376]. Allerdings veränderte KRATZENSTEIN an der Reihenfolge der Bauteile einer *Lingualpfeife* nichts, sondern modifizierte lediglich das Anregungsprinzip.

DUDLEY & TARNOCZY illustrieren sehr treffend die Umstände, unter denen KRATZENSTEINS und KEMPELENS Forschungen stattfanden:

»In the century preceding von Kempelen, there was much speculation by linguists, physicists, psychologists, and speech teachers on the mechanism of speech. The understanding of a physical science was advancing from philosophical speculation

to scientific experimentation with physical apparatus.« [Dudley & Tarcoczy 1950, 157]

Der zweite Teil des »*Tentamen*« schildert, wie die Idee einer »Vokalorgel« praktisch umzusetzen sei (»De construendis fistulis, vocales a, e, i, o, u, enunciantibus«). KRATZENSTEIN thematisiert in jeweils eigenen Abschnitten die Überlegungen, die der Nachahmung der einzelnen Vokale zugrundeliegen. Abschließend rekurriert er noch einmal auf den musikalischen Nutzen seiner Entwicklung. Spätestens im § 31 wird endgültig offenbar, dass KRATZENSTEIN sich der starken Begrenztheit seiner Maschine nur zu bewusst war, aber trotzdem einen halbwegs praktischen Nutzen für sie aufzuzeigen versuchte (vgl. die Übersetzung im Wortlaut im Appendix S. LX).

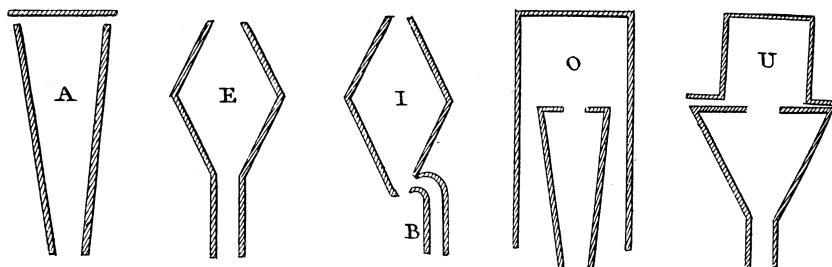
KRATZENSTEIN schwebte demnach eine als Vokalise, also eine nur auf Vokal-Laute gesungene Art der Musik vor. Insbesondere der letzte Satz des Abschnitts entlarvt jedoch die praktische Nutzlosigkeit dieses gut gemeinten, aber doch wohl eher der Verlegenheit entsprungenen Vorschlags zur musikalischen Verwendung seiner »Vokalorgel«: Eine Verteilung der drei genannten Vokale auf die verschiedenen Tonhöhen würde bei strikter Umsetzung des vorgeschlagenen Schemas (c=A, d=E, e=O, fis=A, gis=E, ais/b=O, c=A usw.) dazu führen, dass es gerade einmal drei Töne mit derselben »Vokalqualität« je Oktave gibt, die noch dazu im musikalisch zu jeder Zeit als absolut dissonant empfundenen Intervallverhältnis eines sog. »Tritonus« stehen. Eine von KRATZENSTEIN vorgeschlagene »schlichte Melodie« mit den Tönen derselben Vokalfarbe wäre somit schlicht unmöglich. Hinzu kommt, dass dieses Schema die in der traditionellen europäischen Musik essenziellen Halbtonschritte vollkommen ignoriert.

### 3.2.3 Die »Vokalorgel«

Ergänzend zur schriftlichen Beantwortung der Fragen hatte KRATZENSTEIN eine Vorrichtung zur Vokalsynthese konstruiert, die auf den unter Punkt 2 der Preisfrage formulierten Voraussetzungen basierte. Dem »*Tentamen*« zufolge wurde die Maschine (oder eines von mehreren Exemplaren?) fristgerecht an die Petersburger Akademie geschickt [Kratzenstein 1781, 45 f.].<sup>37</sup> Offenbar wurde der Sendung auch eine Art Bedienungsanleitung beigelegt [Kratzenstein 1781, 39 und Ljubimenko 1937, 346].<sup>38</sup>

<sup>37</sup>Im Brief vom 26. Juli 1780 an J. A. EULER spricht KRATZENSTEIN von einer verbesserten Version, die er nunmehr an die Petersburger Akademie versenden wolle [Ljubimenko 1937, 339].

<sup>38</sup>Diese könnte sich ähnlich der Korrespondenz KRATZENSTEINS im Archiv der Petersburger Akademie befinden.



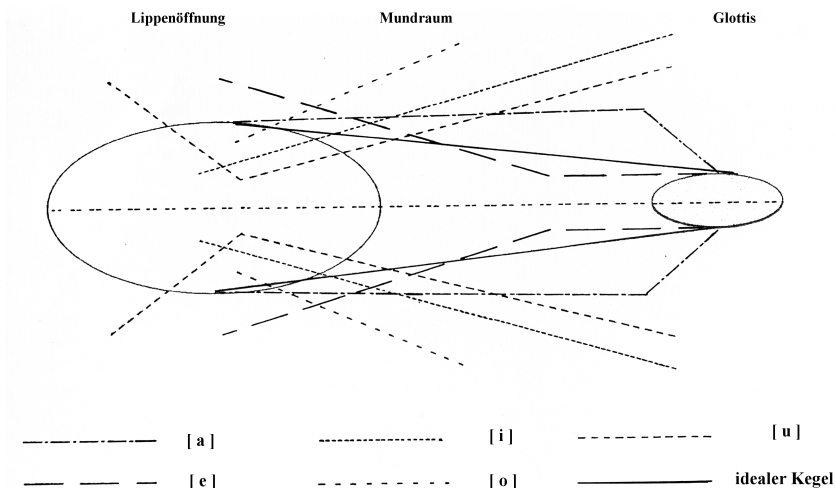
**Abbildung 3.3:** KRATZENSTEINS *Vokalresonatoren* [Young 1845, Pl. XXVI].

Bei der »Vokalorgel« handelt es sich um einen Satz von insgesamt fünf Orgelpfeifen, je eine pro gefordertem Vokal. Ähnlich wie auch bei WOLFGANG VON KEMPELEN (vgl. [Kempelen 1791, 198]) stellte auch für KRATZENSTEIN die Synthese des Lautes I mithilfe einer *Lingualpfeife* ein nicht lösbares Problem dar, sodass er sich stattdessen mit einer Vorrichtung ähnlich einer Blockflöte behalf (vgl. Abb. 3.3). Auch die Umsetzung eines U-ähnlichen Klangs scheint Probleme bereitet zu haben, spricht KRATZENSTEIN im oben erwähnten § 31 doch explizit nur noch von den drei Vokalen A, E und O.

Bezüglich der Entstehung der Vokale hatte er in seiner Schrift die Theorie vertreten, dass der Schall sich ähnlich dem Licht verhalte, woraus er folgte, dass – in Anlehnung an NEWTONS optische Theorie – Resonanzvorgänge analog zu Reflexionen zu behandeln seien. Da es sich beim menschlichen Ansatzrohr nun aber nachweislich um keinen einfachen Trichter handle, müsse seine wechselnde Gestalt durch die Verschachtelung und teilweise Inversion mehrerer Kegel nachmodelliert werden (Abb. 3.4) [Gessinger 1994, 571 f.].

Dieser Annahme folgend konzipierte KRATZENSTEIN für die geforderten fünf Vokale *Resonatoren*, deren Formen er, seiner Theorie zur Vokalbildung folgend, aus sich überschneidenden Kegelschnitten zusammensetzte (Abb. 3.3). Hierbei handelt es sich bei den Resonatoren für A und E um Grundformen, die für die übrigen drei Vokale jeweils abgewandelt wurden.<sup>39</sup>

<sup>39</sup>Diese Grundformen könnten wie das gesamte Anregungskonzept auch, dem Orgelbau entlehnt sein: Für die Gestaltung von *Lingualpfeifen-Resonatoren* finden von jeher sowohl offene oder teilgedeckte Trichter als auch Doppel- und Mehrfachkegel gleichermaßen häufig Verwendung. Die Resonatorformen der »Vokalorgel« für A, E und O sind im Orgelbau des 18. Jahrhunderts vielfältig vorzufinden. Vgl. hierzu bspw. [Bedos 1977, Pl. XVIII und Pl. CXXIX] und [Eberlein 2009 167 f.]



**Abbildung 3.4:** Formen des Ansatzrohrs zur Vokalartikulation in Kegelschnitten nach KRATZENSTEIN (nach Gessinger 1994, 572).

Mit dieser eher nach dem Prinzip des »trial and error« erlangten Konzeption verfehlte er eine an den tatsächlichen anatomischen Gegebenheiten orientierte Synthese:

»With the possible exception of the resonator for the vowel [a] [...], the shapes bear little resemblance to the actual vocal shapes.« [Ohala 2011, 158]

»Die Teile seiner Maschine benennt er nach den Teilen des menschlichen Sprachorgans, welche er für entsprechend hält. [...] Freilich befriedigt diese Nachahmung des menschlichen Ansatzrohrs weit weniger, als die des Kehlkopfs; am besten ist sie noch für a gelungen. Hier erkennt man wenigstens eine Ähnlichkeit mit der Mundform für a. Für die übrigen Klanglaute sind die Formen des Ansatzrohrs lediglich durch Probieren bestimmt.« [Techmer 1884, 92]

GEORG JOSEPH VOGLER attestierte den von KRATZENSTEINS »Vokalorgel« abgeleiteten neuartigen VH-Pfeifen, »den Selbstlauter a gemäß der neuen, von Automaten entlehnten Erfindung ganz deutlich anzugeben« [Vogler o. J., o. S.]. ROBERT WILLIS wies indessen nach, dass die von KRATZENSTEIN unternommenen Anstrengungen mit verschiedenen kom-

plex geformten *Resonatoren* nicht vonnöten waren, sondern dass dieselbe Variation an »Vokalqualität« auch mithilfe einer einzelnen *Zungenpfeife* mit einem schlichten zylindrischen *Resonator* variabler Länge generierbar ist [Willis 1832].

Zur Erzeugung des Sprachschalls griff KRATZENSTEIN vordergründig wie gefordert auf das Prinzip der VH zurück, verwendete also *Zungen-* oder *Lingualpfeifen*. Allerdings übernahm er faktisch – ebenso wie WOLFGANG VON KEMPELEN – nur das grundsätzliche Anregungsprinzip, das *allen* Orgelpfeifen der *lingualen* Bauart zugrunde liegt. Das eigentliche Charakteristikum der VH-Registergruppe, die kurzen, teilgedeckten *Resonatoren*, fand hingegen nur sehr indirekt Anwendung.

KRATZENSTEIN erkannte alsbald die erhebliche Einschränkung, die herkömmliche *Zungenpfeifen* mit sich brachten: Da bei diesen während des Schwingungsvorgangs das metallene *Zungenblatt* auf den Rand der metallenen *Kehle* aufschlägt, entsteht neben dem eigentlichen Ton ein recht starkes metallisches Klirren, das bei der Erzeugung eines vokalähnlichen Tones nicht erwünscht sein konnte und sich auch durch die seinerzeit übliche Methode der *Beledering* der *Kehlenoberseite* nicht völlig unterbinden ließ.

Er umging diese Problematik, indem er *Zungenblatt* und *Kehle* so konstruierte, dass erstere nicht auf den Rändern der letzteren auflag, sondern zwischen ihnen hindurchschwingen konnte. Dieses Funktionsprinzip der *durchschlagenden Zungen* war in Ostasien als Anregungsprinzip der *Mundorgel* bereits seit Jahrhunderten bekannt (vgl. auch Kap. 2.2). Da einzelne Exemplare dieses Instruments zu KRATZENSTEINS Lebzeiten nach Europa gekommen (zumindest eines sogar nach St. Petersburg) und auch in der Literatur beschrieben worden waren, ging man lange Zeit davon aus, er müsse entweder diese Quellen oder aber die Instrumente selbst gekannt und davon beeinflusst seine *Zungenpfeifen* konstruiert haben. Nach AHRENS kann dies aber als widerlegt gelten:

»Zwar erkannte man seit Anfang des 19. Jh. Kratzenstein das Verdienst zu, als erster das Prinzip der Durchschlagzunge in Europa realisiert zu haben, doch unterstellte man, er habe lediglich die Zungen der ostasiatischen Mundorgel kopiert. [...] Überdies seien die von ihm entwickelten Zungen zunächst für die Orgel unbrauchbar gewesen, nicht zuletzt, weil eine wirk-same Stimmvorrichtung fehlte. Beide Annahmen halten einer Überprüfung der Quellen nicht stand. Daß die von Kratzenstein entwickelte Durchschlagzunge einen eigenständigen europäischen Typus repräsentiert, der sich in seiner Konstruktion



(aufgeworfene Zunge auf den Rahmen aufgebunden, aufgenietet oder aufgeschraubt, nicht aber flach in einer Ebene mit diesem liegend wie bei der Mundorgel) und in seinen akustischen Merkmalen deutlich vom ostasiatischen Typus unterscheidet, konnte jüngst nachgewiesen werden. [...] Die Abbildungen im Tentamen belegen zudem, daß Kratzenstein bereits 1780 die Durchschlagzunge mit Stimmkrücken versah.<sup>40</sup> [Ahrens 2003, 617]

Dass das Verdienst der seinerzeit durchaus revolutionären Entwicklung der *durchschlagenden Zungenpfeife* ausschließlich KRATZENSTEIN zukommt, muss bezweifelt werden. Die Konzeption und Fertigung einer *Zungenpfeife* – einer *durchschlagenden* zumal – ist hochkomplex und bedarf umfassender orgelbauerischer Kenntnis. So stellt auch unter anderem AHRENS fest, dass KRATZENSTEIN aller Wahrscheinlichkeit nach eng mit einem versierten Orgelbauer zusammengearbeitet haben muss [Ahrens 2003, 617]. Dem gegenüber steht die Tatsache, dass KRATZENSTEIN als »Instrumentenmacher« (also Werkzeugmacher) einen hervorragenden Ruf besaß.

Ein weiteres Indiz dafür, dass diese Leistung nicht auf einen Alleingang KRATZENSTEINS zurückzuführen ist, liefert der bereits erwähnte ABÉ GEORG JOSEPH VOGLER (1749–1814):

»Der berühmte Professor Kratzenstein in Kopenhagen schickte eine vortreffliche Zeichnung ein [an die Petersburger Akademie der Wissenschaften]<sup>41</sup>, und erhielt die versprochene goldene Medaille; ein sehr geschickter Fortepianobauer [nach Kaufmann 1823, 152: Orgelbauer] in Petersburg, Herr Kirschnick [Franz Kirschnick (1741–1802)], auch ein gebobrner Däne [!], verfertigte nach dieser Zeichnung die Zungen von Messing, und erhielt die silberne Medaille [...].« [Vogler 1810, 128]

Die Lebenswege von KRATZENSTEIN und KIRSCHNICK kreuzen sich auf ganz bemerkenswerte Weise: Während KRATZENSTEIN 1753 von St. Petersburg

<sup>40</sup>In der Tat scheint KRATZENSTEINS Figura 3 eine *durchschlagende Zungenpfeife* mit *Stimmkrücke* (Stimmvorrichtung) darzustellen. Allerdings sprechen die Abbildungslegenden sowohl der lateinischen als auch der französischen Ausgabe die dort abgebildete Pfeife als »gewöhnliche Vox humana« an. Da die Legenden zu den folgenden Figuren 4 und 5 außerdem kontrastierend zwischen »gewöhnlicher Vox humana« und »verbesserte Vox humana« unterscheiden, dürfte es sich bei der Darstellung der Pfeife in Fig. 3 als eine mit *durchschlagendem Zungenblatt* um ein Versehen des Stechers handeln. Schließlich ist das *Zungenblatt* hier so stark verkürzt dargestellt, dass es in dieser Form auch für eine *durchschlagende Zunge* nicht brauchbar wäre.

<sup>41</sup>Aus der zeitgenössischen Korrespondenz ist bekannt, dass Kratzenstein zwar anscheinend *auch* eine Beschreibung der Maschine nach St. Petersburg sandte, vor allem aber die Maschine selbst [Kratzenstein 1781, 45; Ljubimenko, 346].

nach Kopenhagen wechselte und von dort aus 1780/81 die Preisfrage der Petersburger Akademie beantwortete, lebte KIRSnick von 1763–67 in Kopenhagen und zog 1779 oder 1780 nach St. Petersburg [Fredriksson 2002a, 38]. Da KIRSnick dort alsbald mit der Entwicklung der ersten »serienreifen« *durchschlagenden Zungenregister* begann, liegt der Schluss nahe, dass er bereits vorher in Kopenhagen mit dieser Idee in Kontakt gekommen sein muss oder ihm gar ein wesentlicher Anteil an ihr zukommt. Bei KAUFMANN liest sich indessen die Abfolge der Ereignisse anders, als sie gemeinhin dargestellt wird: Demnach hätte zunächst Kirsnick nach dem Vorbild einer *Mundorgel* ein neuartiges *durchschlagendes Zungenregister* entwickelt und KRATZENSTEIN hätte dieses für seine Versuche aufgegriffen [Kaufmann 1823, 152]. Die Frage, wer letztlich auf wessen Erkenntnisse aufbaute, ließe sich wohl nur beantworten, wenn über KIRSnick mehr bekannt würde, als es bislang der Fall ist. Noch ist vollkommen unbekannt, woher er stammte, wo er sich bis in die 1760er Jahre hin aufhielt und wer seine Lehrmeister waren.

Auf den ABBÉ VOGLER ist jedenfalls die rasche Verbreitung der neuartigen *durchschlagenden Zunge* zurückzuführen, da dieser sich nicht nur von KIRSnick's zeitweiligem Mitarbeiter CHRISTOFFER RACKWITZ (1760–1844) 1795 ein Orchestrion mit *durchschlagenden Zungen* bauen ließ, das er auf seinen Reisen quer durch Europa mit sich führte, sondern auch bei mehreren von ihm betreuten Orgelneu- oder Umbauten durch RACKWITZ das neuartige *Register* installierte [Fredriksson 2002a, 40].

Ein reichliches halbes Jahrhundert später urteilte ROBERT WILLIS (1800–1870) über KRATZENSTEIN'S Ansatz, *durchschlagende Zungenpfeifen* zu verwenden:

»Für das Gelingen der Versuche war es wichtig, daß der Ton der Zungenpfeife so weich und rein wie möglich sey. Der rauhe Ton einer gewöhnlichen Rohrpfeife [Lingualpfeife] ist zu diesem Zweck ganz untauglich. [...] [Die Konstruktion einer durchschlagenden Lingualpfeife] war eine sehr wichtige Verbesserung; denn bei sorgfältiger Ausführung dieser Construction erlangte der Ton der Zungenpfeife einen ganz neuen Charakter, und ward der menschlichen Stimme ähnlicher, als der irgend eines bis dahin bekannten Instruments; auch bekam er die nützliche Eigenschaft, daß er, innerhalb gewisser Gränzen, durch stärkeres Anblasen sich verstärken ließ, ohne seine Höhe zu ändern.« [Willis 1832, 402]

Über die Synthesequalität der »Vokalorgel« ist erstaunlich wenig überliefert. Anders als im Falle WOLFGANG VON KEMPELENS scheint es über

KRATZENSTEINS Vokalsynthese keinerlei detaillierten zeitgenössische Berichte zu geben; zumeist begnügt man sich mit dem Hinweis darauf, dass die »Vokalorgel« nur in sehr beschränktem Maße zu Äußerungen fähig war (vgl. [Anonymus 1825, 655]). Als exemplarisch für die Berichterstattung kann hier die Erwähnung in der »*Allgemeinen Deutschen Biographie*« gelten, wo von vier synthetisierbaren Silben die Rede ist (vgl. Zitat auf Seite 51).

Über die synthetisierbaren Silben gibt KRATZENSTEIN selbst Auskunft:

»Pour la douceur de la voix, il faut des tuyaux de carton ou de bois, en place de ceux de metal, ou les revêtir de cuir, de linge ou de carton (on montre dans le Mémoire qui est avec la machine des voyelles, comment il faut disposer cet instrument pour qu'il prononce clairement, papa, mama, nana etc.).« [Kratzenstein 1782, 377]

»Um eine sanfte Stimme zu erhalten, muss man Röhren aus Pappe oder aus Holz statt aus Metall benutzen oder man muss die Röhre mit Leder, Stoff oder Pappe auslegen. (in dem Buch, das der Vokalmaschine beigelegt ist, wird gezeigt, wie das Instrument aufgebaut werden muss, damit es papa, mama, nana usw. klar aussprechen kann).«

Die Übereinstimmung mit den »typischen« Silben bei KEMPELEN ist nicht zufällig, handelt es sich doch um die artikulatorisch simpelsten Lautfolgen. Berücksichtigt man weitere mögliche Silbenfolgen wie »nana«, sind bereits nur unter Verwendung der Pfeife für /a/ drei unterschiedliche Silben synthetisierbar.<sup>42</sup> Falls diese mit den anderen Vokalpfeifen ebenso leicht umsetzbar gewesen sein sollten, verwundert die oben angegebene Gesamtzahl von nur vier unterschiedlichen Silben.

Aus späterer Zeit existieren über die Qualität der Vokalsynthese KRATZENSTEINS widersprüchliche Aussagen:

»Ueber die Einrichtung derselben [der Konstruktion Kratzensteins] ist merkwürdigerweise in den Annalen der Akademie [Petersburger Akademie der Wissenschaften] kein Wort gesagt, woraus sich wohl unschwer der Schluss ziehen lässt, dass Kratzenstein aus seinem Werke, wie viele vor und nach ihm, ein Geheimnis machte. – Tatsache aber ist, dass diese Maschine wirklich die menschlichen Vokaltöne täuschend getreu nachahmen konnte [...].« [Reko 1911, 416]

<sup>42</sup>Leider geht Kratzenstein in seinen Ausführungen mit keinem Wort auf die genaue Funktionsweise seiner »Vokalorgel« ein, so dass unklar ist, wie er den akustischen Unterschied zwischen Plosiv und Nasal bzw. zwischen zwei Nasalen bewerkstelligte.

»I have experimented with cardboard resonators approximating these shapes [die Resonatorformen Kratzensteins] and although I accept that those for [q] and [o] sound more or less like the intended vowel, that for [e] does not. It is possible that what he was aiming for is that the sounds emitted through these resonators at least sounded different.« [Ohala 2011, 158]

Das Schicksal der »Vokalorgel« ist ähnlich ungewiss wie das nahezu aller anderen Sprachmaschinen auch: Allem Anschein nach ist sie rechtzeitig zu einer Präsentation in St. Petersburg eingetroffen, wie die Schlussbemerkungen sowohl der lateinischen als auch der französischen Version des »*Tentamen*« versichern [Kratzenstein 1781, 46; Kratzenstein 1782, 380]. In St. Petersburg verliert sich jedoch ihre Spur, nur wenige Quellen geben mögliche Hinweise darauf, was weiterhin mit ihr geschehen sein könnte. KRATZENSTEIN selbst bestimmt bereits in seinem Brief vom 31. Oktober 1780 an J. A. EULER, dass die Maschine in der Petersburger Akademie verbleiben solle. Pikant wird diese Passage dadurch, dass sich KRATZENSTEIN zugleich erkundigt, wann er mit seiner Prämie für den zuerkannten ersten Preis des Wettbewerbs rechnen könne [Ljubimenko 1937, 347].<sup>43</sup> BREWSTERS Formulierung legt hingegen nahe, dass sich KRATZENSTEIN im Jahre 1786 zumindest mit den Kernteilen seiner Maschine in Paris aufhielt:<sup>44</sup>

»Kratzenstein scheint nicht glücklicher [als Kempelen] gewesen zu seyn, denn ungeachtet er bei seiner Anwesenheit in Paris im Jahre 1786 Herrn de Lalande [wohl Joseph Jérôme Lefrançois de Lalande, 1732–1807, französischer Mathematiker und Astronom] versicherte, er habe eine Maschine angefertigt, welche ganz gut spreche und ungeachtet er ihm den Theil des Apparats zeigte, welcher die Selbstlauter und Worte, wie Papa, Mama aussprach, so läßt sich doch glauben, daß seine Leistungen sich hierauf beschränkten.« [Brewster 1833, 249]

KAUFMANN berichtet:

»Nach dieser Einrichtung verfertigte der Hr. Prof. Kratzenstein in Kopenhagen eine Sprachmaschine, auf der die verschiedenen Körper [Zungenpfeifen] standen und deren Zungen, theils von Messing, theils von Kupfer und theils von Stahl waren, mit

<sup>43</sup>Das Preisgeld dürfte etwa 100 Dukaten (ca. 10.000 € heutiger Kaufkraft) betragen haben [Juškevič & Winter 1959, 11].

<sup>44</sup>Kratzenstein hatte sich bereits 1770 und 1780 in Paris aufgehalten [Fredriksson 2002a, 37 f.].

der er es auch so weit brachte, dass sie die Worte Papa, Mama angab.<sup>45</sup>

[Fußnote:] Eine ähnliche oder vielleicht gar dieselbe befindet sich schon seit vielen Jahren in Paris in den Händen des Hrn. Robertson [Etienne-Gaspard Robert alias Etienne Robertson, 1764–1837, Bühnenmagier und Prof. f. Physik], bey dem man sie täglich sehen kann.<sup>46</sup> [Fußnote Ende]

Hr. Rackwitz, der im Jahre 1791 nach Kopenhagen reisete, um diese Maschine zu sehen zu wollen, bedauert jetzt noch, dass ihm diess nicht möglich wurde, indem man ihm sagte, dass die Maschine nicht gangbar und der Prof. Kratzenstein, schon zu der Zeit sehr bejahrt, krank sey, sich daher nicht sprechen lassen könne.« [Kaufmann 1823, 152]

Es erscheint also möglich, dass sich die Konstruktion 1823 schon seit geraumer Zeit in Paris befand; deshalb wäre im Umfeld des Magiers ROBERTSON weiter nach ihrem Verbleib zu forschen.

Weiterreichende Bedeutung erlangte KRATZENSTEINS »Vokalorgel« erst mit einiger Verzögerung und zudem in einem deutlich anderen als dem ursprünglich intendierten Kontext. Auch allgemein scheint KRATZENSTEINS Arbeit auf dem Gebiet der Sprachforschung und Sprachsynthese im Kontrast zu seinem vielfältigen anderen Arbeiten nur marginale Beachtung gefunden zu haben. So ist den Autoren des KRATZENSTEIN gewidmeten und sehr umfangreichen Artikels in der »Allgemeinen Deutschen Biographie« dieses Thema gerade einmal zwei knappe Sätze wert:

»Als praktischer Mechaniker arbeitete K. an einer Maschine, welche die menschliche Stimme nachahmen sollte und schon vier Silben deutlich wiedergeben konnte. Das Instrument wurde der Petersburger Akademie gesandt.« [Anonymus 1883, 57]

### 3.2.4 Zusammenfassung

CHRISTIAN GOTTLIEB KRATZENSTEIN gehörte zu einer Gruppe junger Wissenschaftler, die von LEONHARD EULER protegiert wurden. Ihnen traute EULER zu, an der Lösung eines Kataloges von ihm als dringlich empfundener naturwissenschaftlichen Fragen zu arbeiten. Vermutlich über seine

<sup>45</sup>Heute wissen wir indessen, dass »Mama« und »Papa« deswegen zumeist die ersten artikulierten kindlichen Äußerungen sind, weil sie die artikulatorisch am einfachsten umzusetzenden sind.

<sup>46</sup>Bei CHAPUIS & GÉLIS ist von einem »Sprechenden Kopf« die Rede, den ROBERTSON im Jahre 1796 in Paris präsentiert habe [Chapuis & Gélis 1928, 210].

hervorragenden Kontakte in der Petersburger Akademie gelang es EULER, eine der Fragen aus diesem Fragenkatalog zur Preisfrage der Akademie des Jahres 1780 zu machen. Seinem Protegé KRATZENSTEIN, der – offenbar außer Konkurrenz – eine umfassende Antwort zu dem Fragenkomplex vorgelegt hatte, wurde daraufhin der erste Preis zugesprochen. Wenn auch bislang handfeste Beweise dafür fehlen, dass es sich hierbei um keinen offenen und fairen Wettbewerb gehandelt hat, legen doch die bekannten Fakten und ähnliche Fälle, bei denen EULER andere Nachwuchswissenschaftler protegiert hatte, diesen Schluss nahe.

KRATZENSTEINS Arbeit stellt einen sicher seltenen Fall in der Wissenschaftsgeschichte dar. Sowohl seine Vokaltheorie als auch die hieraus entwickelte »Vokalorgel« erwiesen sich in wesentlichen Teilen als weitab von jeder Realität. Sein Sprachsynthesemodell war jedoch zwar für seinen originären Zweck wenig tauglich, entwickelte auf gänzlich anderem Gebiet jedoch eine – im Wortsinne – durchschlagende Wirkung. Es muss indessen angezweifelt werden, dass just die Entwicklung eben dieser *durchschlagenden Lingualpfeifen* ausschließlich das Verdienst KRATZENSTEINS war, ungeachtet des hervorragenden Rufs als »Mechanikus«, den er besaß. Vielmehr scheint ein Orgelbauer mit Namen KIRSINICK zumindest mittelbar an der Entwicklung der »Vokalorgel« beteiligt und für die anschließende Etablierung der *durchschlagenden Zungenpfeifen* im Orgelbau verantwortlich gewesen zu sein.

### 3.3 Abbé Mical

»Die biographischen Auskünfte über Abbé Mical, den Urheber der ›Sprechenden Köpfe‹, sind ebenso fragmentarisch und teilweise fragwürdig wie die über das Innenleben seiner Konstruktion.« [Gessinger 1994, 537]

Womit beinahe alles gesagt wäre. NIEMANN zufolge widmete sich der um 1730 geborene MICAL bald nach seiner Priesterweihe ganz seiner Liebhaberei, die der Konstruktion verschiedener humanoider Automaten galt, wie sie auch sein Zeitgenosse JACQUES DE VAUCANSON (1709–1782) erfand [Niemann 1920, 22]. Bereits 1778 hatte er einen ersten sprechenden Kopf konstruiert, der in der Lage gewesen sein soll, einen langen Satz zu sprechen. Zeitgenössischen Berichten zufolge zerstörte MICAL diesen Apparat aus Unzufriedenheit mit seiner Leistung jedoch bald darauf zusammen mit einem »mechanischen Orchester« selbst [Friedrichs 1784, 646]. Ein Neuanfang folgte indessen sogleich (Abb. 3.5). Der französischen Schriftsteller ANTOINE DE RIVAROL (1753–1801) beschreibt in einem *offenen*

*Brief* vom 20. September 1783 MICALS zweite Konstruktion ausführlich [Rivarol 1808, 229 ff.]. Eine gestraffte Übersetzung dieses Briefes findet sich im »*Hannoverschen Magazin*« vom 21. Mai 1784 (im Übrigen gleichfalls wieder als *offener Brief*, nur mit neuem Adressaten und Absender):

»In der Tempelstraße im Marais ist ein mechanisches Werk, welches die Aufmerksamkeit der Kenner dahinzieht, und welches nächstens öffentlich zu sehen seyn wird. Dieses sind zwei Köpfe von Erz, welche reden und ganze Redensarten deutlich und rein aussprechen. Sie sind von colossischer Größe, und ihre Stimme ist übermenschlich; man wird sie nächstens in einem großen Saale aufstellen, damit das Auge und Ohr besser derselben genießen möge.

Es ist dieses nicht, wie Sie leicht denken können, das Werk eines Augenblicks und des Zufalls, wie die aerostatischen Kugeln Montgolfiers, es ist die Frucht der Arbeit und des Genies. Seit dreißig Jahren hat ein gewisser hiesiger Abbé M. Mical mit Vorkehrungen zu diesem glücklichen Ausgange sich beschäftigt; und wenn es möglich wäre alle die Schritte, welche ihn darauf geführt haben, mit dem Auge zu verfolgen, wenn dieser geschickte Künstler alle seine Versuche aufbewahret hätte, so wäre dies eine mechanische Gallerie, deren Besichtigung von äußerstem Nutzen seyn würde. [...]

Nach diesem Grundsatz hat Herr Mical zwei Claviere an seinen redenden Köpfen angebracht: eines in Form eines Cylinders, wodurch man nur eine bestimmte Anzahl Redensarten erhält; aber worauf die Zwischenräume der Worte und ihr Sylbenmaaß genau verzeichnet sind. Das andere Clavier enthält, in dem Umkreise eines Ravalement [Fußnote: Ravalement nennt Herr Mical eine zirkelförmige Figur, deren concave Seite auf gewissen Distanzen kleine perpendiculaire Vertiefungen, die unter einander parallel verlaufen, hat], alle Laute und Töne der französischen Sprache, welche auf eine kleine Anzahl durch eine ingenieuse und dem Verfasser ganz eigene Methode zusammengedrängt sind. Mit ein wenig Fertigkeit und Geschicklichkeit, wird man mit den Händen wie mit der Zunge reden können; und man wird die Sprache der Köpfe die Geschwindigkeit, die Absätze und die ganze Physiognomie, endlich alles das geben können, was eine Sprache hat, die nicht durch Leidenschaften belebt ist.« [Friedrichs 1784, 645 ff.]

Es handelte sich also allem Anschein nach um eine Vorrichtung, die nach dem Prinzip einer mechanischen Orgel aufgebaut war: Über eine *Stiftwalze* konnte ein fest eingestelltes »Programm« abgespielt werden. Was es mit dem *Ravalement* auf sich hat, kann nur gemutmaßt werden. Die im deutschen Text vorfindliche erklärende Fußnote hat keine Entsprechung in RIVAROLS Originaltext. Es scheint sich hierbei um die verunglückte Übersetzung einer möglicherweise ohnehin nicht sonderlich kundigen Erklärung zu handeln. Mit dem ursprünglichen *Ravalement* der Orgel dürfte die Einrichtung jedenfalls nichts zu tun gehabt haben. Vermutlich meinte RIVAROL schlicht eine Erweiterung der zweiten *Walze* gegenüber der erstgenannten, worin diese auch immer bestanden haben mag.

An anderer Stelle mutmaßte RIVAROL, dass eine solche Vorrichtung wie diejenige MICALS dereinst »Taubstummen« zur Kommunikation dienen könnte [Rivarol 1808, 236 f.]. Ebensowenig wie WOLFGANG VON KEMPELEN machte er sich wohl bewusst, dass zur Kontrolle einer wie auch immer gearteten mechanisch vom »Sprecher« gesteuerten akustischen Sprachausgabe zwingend Hörvermögen notwendig ist.

Im Sommer 1783 stellte MICAL, offenbar unter dem Eindruck des Aufenthaltes von KEMPELEN, der seit März in Paris weilte, sich und seine Erfindung der Untersuchung durch eine Kommission der *Académie des Sciences* in Paris, der unter anderem der Mathematiker PIERRE-SIMON LAPLACE (1749–1827), der Chemiker ANTOINE LAURENT DE LAVOISIER (1743–1794) und der Naturforscher CHARLES-GEORGES LE ROY (1723–1789) angehörten. Das hieraus entstandene Gutachten legte am 3. September 1783 das Kommissionsmitglied FÉLIX VICQ D’AZYR (1748–1794) vor. Hierüber berichtet auch LICHTENBERGS »*Magazin*« und bringt eine äußerst knappe Beschreibung der Maschine:

»Der Abbé M\*\*\* hat unterm 2. Jul. d. J. die Pariser Akademie ersucht Commissare zu ernennen, um den Mechanismus zweyer von ihm eingerichteten Köpfe zu untersuchen, welche folgende Reden deutlich aussprechen:

Erster Kopf. Le Roi vient de donner la paix à l’Europe.

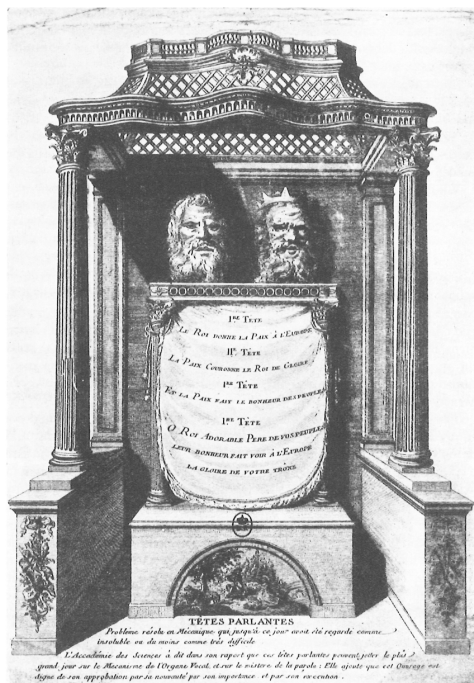
Zweyter Kopf. La paix couronne le Roi de gloire.

Erster Kopf. Et la paix fait le bonheur des peuples.

Hierauf wird die Walze etwas mehr in Bewegung gesetzt, und der erste Kopf wendet sich an den König und ruft aus:

O Roi adorable! Père de vos peuple! Leur bonheur fait voir à l’Europe la gloire de votre Throne. Herr Franklin hat diese Erfindung seines ganzen Beyfalls gewürdigt.« [Lichtenberg 1783, 163]





**Abbildung 3.5:** Idealisierte Darstellung der »Sprechenden Köpfe« des ABBÉ MICAL [Chapuis & Gélis 1928, 205].

KÖSTER referiert ausführlich den Inhalt des Gutachtens, wovon hier nur eine gekürzte Fassung wiedergegeben werden kann, die die wichtigsten Kernpunkte aufzeigt:

»Wenn der Bericht der Akademie auch die umfangreichste Beschreibung der Micalschen Apparatur darstellt, so bleibt er doch unvollkommen und unklar, vor allem fehlen jegliche Abbildungen. Es ist deshalb kaum möglich, über eine grobe Skizze des Mechanismus‘ hinauszukommen. Der Apparat bestand aus zwei Köpfen, die unter Bewegung der Lippen und der Kiefer eine Reihe von Sätzen artikulieren konnten. Im Inneren der Maschine waren – neben einer Windkammer [...] verschiedene als Resonatoren verwendete einfache und doppelte Hohlräume angebracht, die teilweise mit Klappen ausgerüstet waren [...].

Die Resonatoren waren auf der Unterseite durch ein straff gespanntes feines Leder abgeschlossen, das in der Mitte ein ova-

les Loch für die Zungenpfeife, die ihrerseits mit der Luftzufuhr verbunden war, freiließ [...]. Die aufschlagenden Zungen waren durch Metallhaken stimmbar, mit denen die schwingenden Anteile verlängert oder verkürzt wurden. Während der Vorführungen fortlaufender Rede blieb der Grundton jedoch unverändert. Aber schon der Bericht erwähnt die Möglichkeit, die Maschine zum Singen zu bringen. [...]

Wie bei Kratzenstein, so ist auch bei Mical die Anzahl der verwendeten Resonatoren auf fünf (a, e, o, œ, u) begrenzt. Folglich weist der Apparat schon im Vergleich mit dem französischen Vokalinventar große Lücken auf.

Die Öffnungen der Resonatoren sollen den Abmessungen der Kratzensteinischen entsprochen haben. [...] Mit der Maschine konnten neben den genannten Vokalen mit einiger Wahrscheinlichkeit auch die Konsonanten p, b, d, k, g, f, v, l, r erzeugt werden, denn sie erscheinen in dem Text, den die Maschine bei Ausstellungen in der Öffentlichkeit artikuliert haben soll. An ihrer Bildung muß im Wesentlichen das angedeutete Ansatzrohr mit Zunge und Lippen beteiligt gewesen sein.

Ein komplizierter Mechanismus erlaubte die Steuerung koartikulatorischer Effekte. [...]

Das Lauterzeugungsprinzip dieser Micalschen Sprachsynthesemaschine ist dem Kratzensteinischen eng verbunden [...]. Mical war von dem Grundgedanken ausgegangen, daß das stimmgebende Organ in der Glottis nach Art eines mit einer Zunge ausgerüsteten Windinstruments arbeitete, und setzte deshalb [...] als Quelle eine Pfeife mit aufschlagender Zunge ein. [...]

An die Qualität der durch die Apparatur erzeugten Sprache darf man, wie schon bei Kratzenstein, keine allzu hohen Ansprüche stellen. [...]

Da sich die Akademieexpertise nur auf zwei Sätze bezieht, muss man wohl davon ausgehen, daß der Rest nur schwer verständlich war, und selbst diese Sätze sind Gegenstand heftiger Kritik der Akademie. [...] Es ist auch klar die Rede von einer ›imitation très imparfaite de la voix humaine‹. [...]

An der Maschine wurde vor allem kritisiert, daß sie ganze Wörter verschluckte, ihre Stimme rauh und die Artikulation zu langsam sei. [...]

Daß Rivarols Meinung, die Synthesemaschine artikuliere Sätze mit absoluter Klarheit, falsch gewesen sein muß, deutet auch der Umstand an, daß das Publikum die produzierten

Sätze auf einem Plakat ablesen konnte [...]. « [Köster 1973, 82 ff.]

Eine knappe, aber wohlwollende Beschreibung bringt GESSINGER:

»Die ›Musikdose‹ Micals erweist sich also bei näherer Betrachtung als der erste, genauer beschriebene Versuch, die wesentlichen Funktionen des menschlichen Sprechapparates mechanisch nachzubilden. [...] Mical übertraf alle bekannten Konstrukteure sprechender Maschinen im Bestreben, alle Teile des Sprechapparates in ihren Funktionen nachzubilden.« [Gessinger 1994, 538 ff.]

MICAL scheiterte jedoch, ebenso wie KEMPELEN, an der Umsetzung der später auch noch durch HELMHOLTZ vertretenen These DODARTS, dass es sich bei der Glottis um eine membranöse Pfeife handle [Gessinger 1994, 541; vgl. auch Kempelen 1791a, 81 ff. und Tab. II sowie Helmholtz 1863, 160 ff. und Fig. 30.].

Das Fehlen jeglicher exakter Informationen über MICALS »Sprechende Köpfe« scheint auch einer gewissen Scheu, ja Verweigerungshaltung ihres Schöpfers gegenüber dem etablierten Wissenschaftsbetrieb geschuldet zu sein. Er verpflichtete die Kommissionsmitglieder dazu, ihren Bericht nicht ohne seine ausdrückliche Zustimmung zu veröffentlichen. Allerdings starb MICAL offenbar auch bereits um 1790<sup>47</sup>, so dass die Entwicklung jäh abbrach [Gessinger 1994, 538].

Abschließend seien zwei Aspekte hier als bemerkenswert herausgestellt: Zum einen ist die »Sukzession« gewahrt; die Sprechmaschine (nicht Sprachmaschine!) des ABBÉ MICAL präsentiert sich, noch Jahrhunderte nach dem Philosophen und legendenhaften Pionier der Sprachsynthese ROGER BACON, als bronzenen Kopf (bzw. Köpfe). Dies dürfte wohl kaum ein Zufall gewesen sein. Zum anderen macht eine genauere Betrachtung der zeitlichen Zusammenhänge deutlich, dass im Jahr 1783 Paris so etwas wie der Kumulationspunkt der mechanischen Sprachsynthese wurde: Im Jahr zuvor war KRATZENSTEINS »*Tentamen*« in Paris in der französischen Übersetzung erschienen, im Jahr selbst hielt sich KEMPELEN mit »Schachtürken« und Sprachmaschine dort auf und MICAL ließ – sicher auch durch diesem Umstand motiviert – seine eigene Sprachsynthese durch die Pariser Akademie examinieren. Und just in diesem Jahr starb der geistige Vater der seinerzeit modernen Sprachmaschinen, LEONHARD EULER.

<sup>47</sup>WILLIS nennt ohne Quelle das Jahr 1786 [Willis 1832, 398], KÖSTER und HANKINS & SILVERMAN geben jeweils 1789 an, ebenfalls ohne Quellenangabe [Köster 1977, 81, Hankins & Silverman 1995, 188].

### 3.4 Wolfgang von Kempelen

»Kempelen's Werk ist reich an originellen und glücklichen Gedanken, eben so merkwürdig durch den Scharfsinn, den der Verfasser darin entwickelt, wie durch die in der That lebendige und unterhaltende Art der Behandlung des Gegenstandes.«  
[Willis 1832, 399]

#### 3.4.1 Historische Aspekte der Sprachmaschine

Um die Person und das Wirken WOLFGANG VON KEMPELENS<sup>48</sup> (1734–1804) hat sich so stark wie um keinen seiner »Kollegen« ein schier undurchdringliches Gewirr an Legenden, Halbwahrheiten und fehlerhaften Zuschreibungen gebildet. Das Verdienst, auf der Grundlage intensiver und solider Recherche hier erstmals grundlegend und nachhaltig für Klarheit und Entzauberung gesorgt zu haben, kommt ohne Zweifel ALICE REINIGER zu [Reiniger 2007]. Für nähere Angaben zur Person KEMPELENS sei daher auf ihre Arbeit verwiesen, so dass hier eine Beschränkung auf die notwendigen Fakten möglich ist. Eine detaillierte und kritische Auseinandersetzung mit der Sprachmaschine findet sich bei BRACKHANE (2011).

WOLFGANG VON KEMPELEN wirkte ab 1755 in verschiedenen Funktionen in der Administration des k. k. österreichischen Staats. Neben seinen vielfältigen (stets zur vollen Zufriedenheit Kaiserin MARIA THERESIAS (1717–1780) erfüllten) Aufgaben, die allein ihn bereits als einen vielfältig interessierten und findigen Kopf kennzeichnen, konnte KEMPELEN sich im Laufe mehrerer Jahrzehnte einen Ruf als äußerst geschickter Ingenieur im Wortsinne erwerben. In der Reihe seiner Erfindungen darf der berühmterberüchtigt gewordene »mechanische Schachspieler« oder »Schachtürke« von 1769 ff., der seinerzeit einen lebhaften Widerhall in der Berichterstattung fand, als sein Opus magnum gelten, dicht gefolgt von der Sprachmaschine. Während die Spielstrategie und mögliche Funktionsweise des »Schachtürken« in einer Vielzahl zeitgenössischer Zeitungsberichte, Essays und ganzen Büchern intensiv aus philosophischer und mathematischer Sichtweise abgehandelt wurde, fand die Sprachmaschine stets nur am Rande Erwähnung. Dies ist vermutlich auch dadurch begründet, dass sie just zu dem Zeitpunkt der Öffentlichkeit präsentiert wurde, als durch die Publikationen KARL GOTTLIEB WINDISCHS (1725–1793) das Rätsel um den angeblich androiden »Schachtürken« zumindest in Teilen gelöst wurde und sich eine gewisse Ernüchterung über die »Trivialität« und den scheinbaren

<sup>48</sup>In zeitgenössischen Quellen häufig »Kempele« oder »Kempel« genannt.

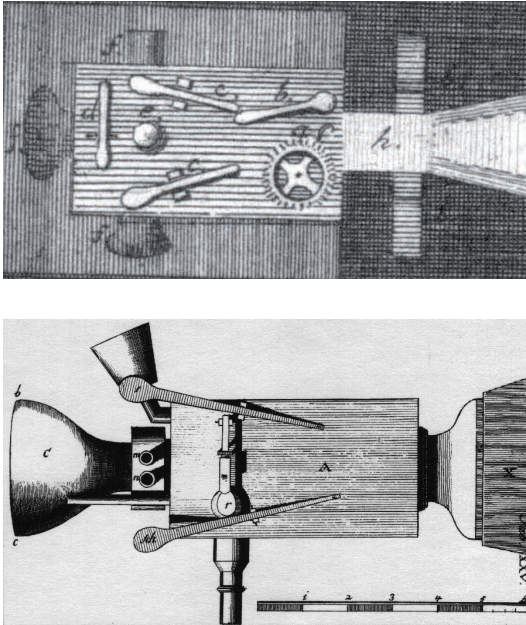
Betrug am Publikum einstellte. Nachdem sich der »mechanische Schachspieler« als gehobener Taschenspielertrick erweisen hatte (KEMPELEN seinerseits hatte nie etwas Gegenteiliges behauptet), traute man auch der Sprachmaschine nicht mehr zu.

KEMPELENS Beschäftigung mit den sog. »Taubstummen« und einer möglichen Hilfestellung zur Kommunikation für diese fiel in eine Zeit, in der auf diesem Gebiet sowohl in philosophischer als auch in praktischer Hinsicht einiges in Bewegung geriet. Bis in KEMPELENS Zeit hinein wogte die – aus heutiger Sicht wohl absurd erscheinende – Diskussion darüber, ob Menschen, denen die Fähigkeit zu hören und zu sprechen fehle, überhaupt als Menschen im engeren Sinne zu betrachten seien, oder ob man es nicht vielmehr entweder mit vernunftbegabten Tieren oder aber mit Maschinen in Menschengestalt zu tun habe. Entsprechende Argumentationen vertraten JOHANN CHRISTOPH GOTTSCHED (1700–1766), JULIEN OFFRAY DE LA METTRIE (1709–1751) oder KARL PHILIPP MORITZ (1756–1793), die einen direkten Zusammenhang zwischen Vernunftbegabung, Sprachfähigkeit und Denkkraft herstellten. KEMPELEN hingegen hielt es mit DESCARTES (1596–1650) und DIDEROT (1713–1784): Die lediglich organische Verhinderung von Sprache mache einen Menschen nicht zum Tier, da die Fähigkeit hierzu eine Begabung durch Vernunft sei [Kempelen 1791a, 1]. Maßgeblich beeinflusst waren diese Ansichten wohl durch seine Kontakte zu den Taubstummenlehrern CHARLES-MICHEL DE L'ÉPÉE (1712–1789) in Paris und JOHANN FRIEDRICH STORK (1746–1823) in Wien (vgl. KEMPELENS eigene Angaben [Kempelen 1791a, 17 ff.]).<sup>49</sup>

KEMPELENS erste Überlegungen – zunächst nur zur Vokalsynthese – entstanden vor 1769 [Kempelen 1791a, 388 ff.]. Bis zur Veröffentlichung des »*Mechanismus der menschlichen Sprache*« [Kempelen 1791a, b] mit Präsentation eines zumindest in wesentlichen Teilen gefestigten Sprachsynthesekonzeptes sollten von da an noch über zwei Jahrzehnte vergehen.

Spätestens 1783 existierte jedoch ein bereits grundsätzlich vorzeigbarer Prototyp, den KEMPELEN zusammen mit dem »Schachtürken« auf eine regelrechte Europatournee mitnahm, die ihn unter anderem nach Paris und London führte. GESSINGER zufolge hielt er sich jedoch bereits im Jahr zuvor in St. Petersburg auf, wo er den »Schachtürken« vorführte und mutmaßlich auch KRATZENSTEINS Vokalsynthese kennenlernte [Gessinger 1994, 589]. Diese »Tournéeversion« der Sprachmaschine unterschied sich jedoch in ihrer Komplexität und Bedienung merklich von der späterhin im »*Mechanismus*« gezeigten Form, wie die Beschreibung und insbesondere die detaillierte Abbildung bei HINDENBURG zeigt [Hindenburg 1784] (vgl. Abb.

<sup>49</sup> Offenbar war KEMPELEN sogar an Planungen zur Einrichtung einer neuen Institution nach DE L'ÉPÉES Vorbild beteiligt [Schumann 1940, 199].



**Abbildung 3.6:** Die Sprachmaschine KEMPELENS bei [Hindenburg 1784] (oben) und [Kempelen 1791a, Tab. XXV] (unten).

3.6). Da dieser explizit betont, dass der Stich nicht nur durch KEMPELEN autorisiert, sondern auch redigiert worden sei, kann auf eine große Authentizität der Darstellung geschlossen werden [Hindenburg 1784, 47].

Wohl im Zusammenhang mit KEMPELENS Aufenthalt in Paris zwischen März und Herbst 1783 und der damit verbundenen Berichterstattung ließ der ABBÉ MICAL seine Maschine im Juli 1783 durch die Pariser Akademie begutachten (vgl. Kapitel 3.3). RIVAROL kommentiert maliziös:

»M. Kempelein [sic!] avoit aussi un coffret d’ou il s’échappoit quelques mots, à ce qu’on dit: Mais cette honnête Voyageur a rendu un véritable hommage à M. le Abbé Mical; des qu’il a eu connoissance des Têtes-parlantes, il à retiré son Automate, son coffret & sa propre personne.« [Rivarol 1808, 243]

»Herr Kempelen besitzt ebenfalls einen Kasten, dem verschiedene Äußerungen entchlüpfen, wie man sagt: dieser ehrbare Reisende hat den Herrn Abbé Mical aufrichtig gewürdigt; nachdem er über dessen sprechen-

de Köpfe Kenntnis erlangt hat, hat er seinen Automaten [den »Schachspieler«], seinen Kasten und sich selbst zurückgezogen.«

Die wohl lebhafteste Berichterstattung und Fürsprache findet KEMPELEN ab 1784 in LICHTENBERGS »*Magazin für das Neueste aus der Physik und Naturgeschichte*«. Beschränkt sich der erste Artikel noch auf eine verhältnismäßig kurze Notiz hinsichtlich der Anfeindungen, denen KEMPELEN offenbar 1783 auf seiner Tournee in Frankreich und England ausgesetzt war [Lichtenberg 1784, 219 f.], so bricht LICHTENBERG im Jahr darauf eine Lanze für ihn, unter anderem durch den Hinweis, dass gerade die Imperfektion und die Fähigkeit der Maschine, sowohl zu flüstern als auch laut zu sprechen, deutlich gegen einen auf Bauchrednerei basierenden Betrug sprächen [Lichtenberg 1785, 183 ff.]. Nach dem Erscheinen des »*Mechanismus*« 1791 veröffentlichte LICHTENBERG dann 1792 eine umfangreiche Paraphrase der »*V. Abtheilung – Von der Sprachmaschine*« und zusätzlich eine ebenso umfassende wie wohlwollende Rezension des gesamten Buches [Anonymus 1792a, Anonymus 1792b, 127 ff.].

Nachdem die Vorführungen insbesondere der Sprachmaschine in London und Paris nicht auf die erwartete Resonanz stießen und er sich im Gegenteil in mehreren Periodika der Jahre 1784/85 mit handfesten Polemiken konfrontiert sah, erkannte KEMPELEN wohl spätestens jetzt die Notwendigkeit, seine Synthese-Konstruktion offenzulegen. So war beispielsweise 1784 in der »*Berlinischen Monatsschrift*« zu lesen:

»Noch mehr wäre eine nähere Anzeige die Pflicht des Herrn v. Kempelen bei der Redemaschine. Da, wie jeder Vernünftige weiß, alles was man bisher in der Art gezeigt hat, grober Betrug war; und da mehrere Kenner behaupten, dass es unmöglich sei, artikulierte Töne von einer Maschine hervorbringen zu lassen; so müßte Herr von Kempelen augenscheinlich und handgreiflich zeigen, daß und wie sein Schlagen auf den Tangenten [Spielen auf den Tasten der Maschine] diese artikulierten Töne hervorbringe. Bis dahin thut man wohl zum Besten, davon zu schweigen.« [Biester 1784, 514]

Eine recht ausführliche, aus heutiger Sicht aber abstruse Polemik bringt der unter dem Kürzel RM schreibende Autor, der in der »*Allgemeinen deutschen Bibliothek*« die Langfassung der »*Briefe über den Schachspieler*« KARL GOTTLIEB WINDISCHS rezensiert, ohne die Sprachmaschine selbst je zu Gesicht bekommen zu haben:

»Indessen, der Briefverfasser [Karl Gottlob Windisch] giebt uns schon S. 46 ff. [In [Windisch 1783a]. In [Windisch 1783b]: 25

ff.] von einer neuen Kempelschen Maschine Nachricht, welche spricht. Er scheint sich ganz zu vergessen, indem er hinzusetzt: ›Gestehen Sie doch, man muß schon einen schöpferischen Geist, ein unbändig kühnes Genie haben, wenn man sich an so etwas wagen darf, und es geräth ihm schon wieder. Er ist schon so weit damit gekommen, daß er die Möglichkeit einer solchen Maschine zeigen und den Gelehrten eine neue und bisher unbekannte Erfindung zu beurtheilen vorstellen kann... Schon beantwortet sie einige Fragen ziemlich deutlich und vernehmlich; ihre Stimme ist ein sanfter Alt, der Ton auch ganz angenehm, das R jedoch spricht sie etwas schnarrend aus. Wenn man etwas nicht versteht oder verstehen will: so wiederholt sie das Gesagte langsam.‹ Also kann die Maschine hören und verstehn, was man sagt! Man sollte mit solchen Inkonsequenzen kaum Mitleiden haben. ›Wenn man dieses aber noch einmal fordert, so sagt sie es mit einer bösen und aufgebrachten Stimme.‹ Also hat sie auch Leidenschaften! Was Hr. von Kempelen nicht schaffen kann! ›Ich habe sie in verschiedenen Sprachen Wörter und Redensarten ganz und vernehmlich aussprechen gehört, die ich hier mittheilen will z. B. Papa ff.‹ Ich übergehe diese Wörter, um die Leser nicht gar zu sehr zu äffen. Ich bemerke nur, daß es größtentheils Lippen und Kehlbuchstaben sind, woraus diese Wörter bestehn; und dieß veranlaßt mich zu schliessen, daß das Sprachrohr, durch welches der versteckte Mensch aus der Maschine spricht, noch nicht diejenige Richtung erhalten hat, welche erforderlich ist, um Zahn- Gaum- Zungen- und Nasenbuchstaben durch den Mund der Maschine zu stoßen [...]. Von der Schachmaschine wird nun endlich bekannt, daß sie ein Blendwerk ist, nachdem man sie viele Jahre für eine wirklich spielende Figur gehalten hat. Schon deshalb kann niemand Vertrauen auf die sprechende Maschine setzen. Aber überhaupt muß man erstaunlich unwissend seyn im Bau jedes Sprachorgans und seiner Widmung, und im Zusammenwirken sämtlicher Organen zur Artikulation menschlicher Stimme, kurz im Mechanismo menschlicher Sprache, wenn mans selbst glauben und andere überreden will, daß es nur im geringsten möglich sey, Maschinen zu erfinden, welche artikulierte Töne zu reden vermögen. [...] Aber ich wünschte, daß der V. [Verfasser] und auch der H. von Kempelen sich mit den Schriften eines Paul Bonnet, [...], Helmont [...] bekannt machen möchten, um



daraus das Gebäude der Menschengesprache kennen zu lernen«  
[Rm 1784, 278 ff.]

FRANCISCUS MERCURIUS VAN HELMONT (1614-1699) hatte in der Mitte des 17. Jahrhunderts eine Theorie entwickelt, nach der das Hebräische die Ursprache des Menschen sei, da jeder hebräische Buchstabe »das Abbild der Bewegungen der Zunge« [Schumann 1940, 74] bei der Aussprache des repräsentierten Lautes darstelle [Helmont 1657]. Nachdem bereits KEMPELEN diese Theorie ebenso grundsätzlich wie handfest verworfen hatte [Kempelen 1791a, 143 ff.], urteilt PAUL SCHUMANN 1940:

»Der Nachweis dieser Übereinstimmung führt Helmont zu vielen Gewaltsamkeiten, und auch die von J. Franck entworfenen Abbildungen, die hebräische Buchstaben mit den entsprechenden Sprechbewegungen des Menschen (die innere Sprechpartie ist sichtbar gemacht [d. i. Sagittalschnitte]) darstellen und durch mehrere Zungen und deren verschiedene Stellung auch die fortlaufenden Bewegungen andeuten wollen, sind übertrieben und unnatürlich.« [Schumann 1940, 74]

Die Beobachtung, dass KEMPELENS Sprachmaschine insbesondere »Lippen und Kehlbuchstaben« spricht, ist indessen vollkommen korrekt und hauptsächlich dem Fehlen eines Zungenanalogons geschuldet. Die von RM daraus gezogenen Schlüsse sind es jedoch nicht. Allerdings muss dem Rezensenten zugestanden werden, dass die recht großsprecherische und stark an eine Werbeschrift erinnernde Stilistik WINDISCHS nicht gerade dazu angetan war, kritischen Geistern Vertrauen in die so hochgepriesenen Maschinen zu verleihen.

Auch FRIEDRICH NICOLAI (1733-1811), prominenter Hauptvertreter der *Berliner Aufklärung*, meldete sich mit substanzieller Skepsis zu Wort, die aber zugleich auch zeigt, wie grundsätzlich die Funktionsweise der Sprachmaschine von ihm und anderen Skeptikern (vgl. RM) missverstanden wurde:

»Was man von der Maschine des Hr. v. Kempelen, welche deutsch, lateinisch, französisch und italienisch sprechen soll, erwarten könne, ob sie Mechanismus oder Täuschung sey, ist leicht zu erachten. Herr P. Hindenburg hat sie abbilden lassen. Man sieht hier zwar auch Hebel und Räder, aber man kann von einem vorgeblichen Mechanismus, welcher artikulierte Töne hervorbringen könnte, nicht das geringste sehen. [...] Ein rußischer Künstler Kischnick [der Petersburger Orgelbauer FRANZ KIRSCHNICK (1741-1802), der nach dem Vorbild der

KRATZENSTEINSCHEN Vokalpfeifen Orgelregister entwickelte, vgl. Kap. 3.2.3] sollte drei Vokale [!] herausgebracht haben; das hielt man schon für viel. Aber doch konnte es Hr. Bernoulli nicht zu hören bekommen. [...] Und der Hr. v. Kempelen will Vokale und Konsonanten und Wörter heraus bringen? Seine Sprachmaschine giebt einen feinen kindischen Ton von sich; es ist also wohl zu erachten, daß eben das Kind, welches im schachspielenden Türken den Arm bewegt, auch hier seine kindische Stimme vermittelt einer sehr bekannten akustischen Vorrichtung hören läßt. Es kann einer der Füße des Tisches, worauf dieß Kästchen steht, hol seyn, und auf ein kleines Loch im Boden passen. Der Schall kann auf mannichfaltige andere Art hineingebracht werden. Um zu entdecken, wie es geschieht, ist nicht bloß ein Mathematiker nöthig wie Herr Hindenburg, der ohne Betrug zu vermuthen, nur bloß wissenschaftliche Theorie sucht.« [Nicolai 1785, 430 ff.]

Vor dem Hintergrund dieser fundamentalen Skepsis erklärt sich die auch von GESSINGER richtig erkannte Tatsache, dass der »*Mechanismus*« deutlich erkennbar von hinten aufgekläumt, also vollständig im Hinblick auf die Sprachmaschine konzipiert wurde und in deren Beschreibung kulminiert [Gessinger 1994, 583 f.]. Im Übrigen mutet KEMPELENS Ansatz, im »*Mechanismus*« die Konstruktion seiner Sprachmaschine zum Zwecke der Weiterentwicklung durch die interessierte Öffentlichkeit weitgehend offenzulegen, geradezu wie eine Frühform des »Open-Source«-Gedankens an.<sup>50</sup>

Auch Jahrzehnte nach dieser Publikation findet sich noch ein in doppeltem Sinne konstruktiver Hinweis zur Bauweise der Sprachmaschine, der zugleich zeigt, wie stark KEMPELENS Ruf allem Anschein nach war und wie wenig Zweifel man nunmehr an der grundsätzlichen Richtigkeit seines Ansatzes hatte. Der Naturforscher DAVID BREWSTER (1781–1868) schreibt:

»Hätte Herr von Kempele die neuere Entdeckung gekannt, dem Caoutschuk jeden Grad von Weiche zu geben, indem man es mit Melasse oder Zucker vermischt, welche Verbindung stets Feuchtigkeit aus der Atmosphäre absorbiert, so wäre seine Nachahmung der menschlichen Organe noch vollkommener gewesen.« [Brewster 1833, 246, Fußnote]

<sup>50</sup> Allerdings präsentierte Kempelen die Sprachmaschine grundsätzlich nur mit geschlossenem Außenkasten, dessen Öffnungen er zusätzlich mit Tüchern verhängt hatte. Ob dies der grundsätzlichen Geheimhaltung oder eher der Dramaturgie seiner Vorführungen diene, muss offen bleiben.

Laut BREWSTER erhielt man, wenn man bei KEMPELENS Maschine nur ein »Nasenloch« verschloss, ein U und nicht wie bei Kempelen geschildert ein N. Möglicherweise liegt hier jedoch schlicht ein Druckfehler vor. Interessant ist auch der Hinweis, dass Kempelen seine Sprachmaschine deshalb nur selten öffentlich gezeigt habe, weil er mit ihrer effektiven Leistung unzufrieden gewesen sei [Brewster 1833, 247].

Der Übersetzer von BREWSTERS »*Briefen*«, der Berliner Naturwissenschaftler FRIEDRICH WOLFF (1766–1845) beschreibt in einem raumgreifenden Fußapparat zu BREWSTERS Ausführungen seine eigenen Erlebnisse mit KEMPELENS Sprachmaschine:

»Herr von Kempele hatte die Gefälligkeit, mir seine Sprachmaschine in Leipzig im Jahre 1785 [recte: 1783] zu zeigen. [...] Das Sprechen war so täuschend, daß, wenn man das Gesicht abwandte, unwillkürlich das Auge dem Ohre folgte, und man das Gesicht der Maschine zukehrte. Auf meine Frage, wie es ihm geglückt wäre, alle bei der Construktion einer solchen Maschine sich darbietenden Schwierigkeiten zu überwinden? Erwiederte Herr von Kempele: In den Tönen, welche mehrere Thiere von sich geben, kämen die Laute gewisser Buchstaben besonders deutlich vor, z. B. das M im Brüllen einer Kuh, das B im Blöken des Schafes, das R im Gekrächze des Rabens u. s. w. Durch genaue Untersuchung der Stimmritzen der verschiedenen Thiere, und durch ähnliche Einrichtung der Lippen in den Pfeifen, wäre er, nach unzähligen Versuchen, mit seinem Kunstwerk zu Stande gekommen.« [Brewster 1833, 248, Fußnote]

Da sich von diesem sehr eigenartigen deduktiven Vorgehen in KEMPELENS späteren schriftlichen Ausführungen nicht auch nur der geringste Hinweis mehr findet, bleiben zwei denkbare Möglichkeiten, wie es zu dieser bemerkenswerten Äußerung gekommen sein kann: Entweder war KEMPELEN zu jener Zeit in seinen anatomisch-artikulatorischen Ansichten tatsächlich noch eine beträchtliche Wegstrecke von seinen späteren Erkenntnissen entfernt, oder aber – was wahrscheinlicher sein dürfte – FRIEDRICH WOLFF nahm diese ironische Bemerkung KEMPELENS irrtümlich für bare Münze.

Allerdings ist in der Tat davon auszugehen, dass KEMPELEN, anders als von ihm selbst dargestellt, die intensive Auseinandersetzung mit den Schriften zur Sprachproduktion und Physiologie nicht vollständig am Beginn seiner Forschung vornahm, sondern parallel oder gar erst im Nachgang zur Arbeit an der Sprachmaschine, nämlich gezielt in Hinblick auf die Publikation des »*Mechanismus*«, um diesem ein stabiles theoretisches Fundament zu verleihen. Hierfür spricht, dass die von KEMPELEN in sei-

**Tabelle 3.2:** Chronologische Erscheinungsreihenfolge der von KEMPELEN referenzierten Grundlagenliteratur zum »Mechanismus«.

Autor	Kurztitel	Jahr
Wilkins, John	» <i>An Essay Towards a real character...</i> «	1668
Helmont, Franciscus M. v.	» <i>Alphabeti vere naturalis...</i> «	1676
Dodart, Denis	» <i>Mémoire sur les causes...</i> «	1703
Ferrein, Antoine	» <i>De la formation de la voix de l'homme</i> «	1741
Brosses, Charles de	» <i>Traité de la formation mécanique...</i> «	1765
Haller, Johann A. von	» <i>Elementa physiologicae...</i> « Bd 3	1766
Süßmilch, Johann P.	» <i>Versuch eines Beweises...</i> «	1766
Herder, Johann G. v.	» <i>Abhandlung über den Ursprung der Sprache</i> «	1772
Kalmar, Jiří	» <i>Praecepta grammatica...</i> «	1772
Zobel, Rudolf H.	» <i>Gedanken über die Meinungen der Gelehrten...</i> «	1773
Court de Gébelin, Antoine	» <i>Monde primitif...</i> «	1775
Lieutaud, Joseph	» <i>Zergliederungskunst...</i> «	1782
D'Azyr, Vicq	» <i>Mémoire sur la voix...</i> «	1782

nem Buch angeführten Werke zum Teil erst deutlich nach 1769 erschienen (vgl. Tab. 3.2).

Rätselhaft ist der Umstand, dass KEMPELEN wider besseres Wissen eine »handelsübliche« *Zungenpfeife aufschlagender* Faktur benutzte. Im »*Mechanismus*« stellte er selbst fest, dass die Pfeifen der VH für seine Zwecke wenig geeignet seien [Kempelen 1791a, 390 ff.]. Zudem muss es als gesichert gelten, dass er recht genau über die Vokalsynthese KRATZENSTEINS informiert war [Kempelen 1791a, 197 ff.], sie bei seinem Aufenthalt in St. Petersburg 1782 vermutlich sogar gesehen und gehört hatte [Gessinger 1994, 589].

KEMPELENS Wahl des Anregungskonzepts könnte indessen eine Verlegenheitslösung gewesen sein: Er hatte – nicht als erster und angeregt durch die Überlegungen DODARTS – ursprünglich die Idee, als Anregungsinstrument eine *membranöse Pfeife* zu verwenden [Kempelen 1791a, 80 ff. und Tab. II].<sup>51</sup>

<sup>51</sup>Siehe Kap. 8.2.6

Zu dem in der zeitgenössischen Berichterstattung immer wieder auftauchenden Assistenten bzw. Vertreter Kempelens bei Vorführungen insbesondere des »Schachtürken«, einem Herrn ANTHON, liefert GESSINGER (leider ohne Nennung seiner Quelle) einen spekulativen Hinweis:

»Bei diesem ständigen Begleiter der Schachmaschine könnte es sich um den Wiener Mathematiklehrer Anton Ferkel handeln, der 1780 einen Sprachsyntheseapparat gebaut haben soll.« [Gessinger 1994, 420, Fußnote 8]

Angeblich sprach ausgerechnet LEONHARD EULER KEMPELEN die Befähigung ab, eine Sprachmaschine zu bauen, da dies wider die göttlichen Gesetze sei [Reko 1911, 416].<sup>52</sup> Auf die zeitgenössische Berichterstattung über KEMPELENS Konstruktion wurde bereits in BRACKHANE (2011) ausführlich eingegangen. Die retrospektive Bewertung von KEMPELENS Arbeit insbesondere im 19. Jahrhundert stützte sich – mangels Zeugnissen aus erster Hand – auf diese Berichte und kolportierte sie – oftmals legendenhaft angereichert – weiter. Ungewöhnlich kritisch heißt es hingegen in der Neuauflage von »Gehler's *physikalischem Wörterbuch*«:

»Der Process [des Sprechens] ist ein sehr zusammengesetzter und durch künstliche Mechanismen keineswegs im gleichen Grade nachbildbar, als dieses bei der Gesangsstimme der Fall ist, und v. Kempelen's Sprachmaschine giebt bei einiger Nachhülfe mit der Hand kaum den menschlichen entfernt ähnliche Laute.« [Gehler 1836, 387]

Dieses Zitat ist insbesondere daher bemerkenswert, als es suggeriert, dass der Autor sich selbst einen Eindruck von der Synthesequalität der Maschine verschaffen konnte. Da sich in der Ursprungsausgabe des Wörterbuches von 1787 zwar eine Notiz über KEMPELENS »Schachspieler« findet, nicht jedoch über die Sprachmaschine, so ist davon auszugehen, dass die Beurteilung letzterer auf einer eigenen Beobachtung der Bearbeiter von 1825 ff. zurückzuführen ist.

Von KEMPELENS Maschine fehlt indessen spätestens seit seinem Tode jede Spur. Es ist zu vermuten, dass sie zusammen mit dem »Schachtürken« in den Besitz JOHANN NEPOMUK MÄLZELS (1772–1838) kam [Pompino-Marschall 1991, 225] (siehe auch 4.4.1, S. 98). Ob sie jedoch dessen Tourneen insbesondere nach Amerika mitmachte, ist bislang völlig ungesichert. Überliefert ist, dass der »Schachtürke« unter MÄLZELS Ägide in der Lage

<sup>52</sup>Hierzu passen indes die im Abschnitt 3.2.2 angeführten Zitate EULERS (S. 33), aus denen seine Ansicht, die Sprache sei eine Gottesgabe, deutlich hervorgeht.

war, Äußerungen wie etwa »Schach« von sich zu geben und dass MÄLZEL auch selbst eine »Mama« und »Papa« sagende Maschine gebaut hatte, die große Ähnlichkeit mit derjenigen Kempelens hatte [Hankins & Silverman 1995, 213].<sup>53</sup> Ob es sich hierbei jedoch um dessen Sprachmaschine handelte oder um eine für die Zwecke des »Schachtürken« adaptierte neue, ist unklar. Jegliche Nachforschungen sind indessen zum Scheitern verurteilt: Als KEMPELENS Sohn 1822 starb, wurde zumindest der gesamte Familienbesitz in Gomba (Ungarn) verkauft. Was damals nicht in andere Hände geriet, kam der Familie nach und nach im Laufe der nächsten Jahrzehnte abhanden, bis schließlich der Zweite Weltkrieg auch noch die Vernichtung des Familienarchivs besorgte [Reininger 2007, 28].

### 3.4.2 Repliken von ca. 1800 bis heute

Glaubt man der eher beiläufigen Anmerkung des KEMPELEN und seiner Sprachmaschine höchst kritisch gegenüberstehenden Autors RM, so gab es bereits 1784 (!) erste Versuche, KEMPELENS Maschine nachzubauen, sieben Jahre vor dem Erscheinen des »*Mechanismus*« [Rm 1784, 280]. Dabei handelte es sich aber allem Anschein nach um eine anthropomorphe Figur, bei der nach Auskunft RMS die Stimme aus der Bauchgegend erscholl. Nach Veröffentlichung des »*Mechanismus*« scheint es dann sehr rasch erste seriöse Nachbauten von KEMPELENS Sprachsynthesekonstruktion gegeben zu haben. Dies ist nicht zuletzt dem Umstand geschuldet, dass nunmehr erstmalig eine tatsächlich detaillierte »Bauanleitung« vorlag, verbunden mit der ausdrücklichen Aufforderung des Autors, sein Werk zur Vervollendung zu bringen [Kempelen 1791a, 456]. Zumindest einige dieser Repliken fanden ihrerseits wieder öffentliche Beachtung. So erwähnt beispielsweise JOHANN WOLFGANG VON GOETHE (1749–1832) 1797 einen Nachbau, dessen akustische Charakterisierung mit derjenigen des Originals in der zeitgenössischen Presse konform geht [Goethe 1893, 154]. Bis zum heutigen Tage entstand eine ganze Reihe von Nachbauten, die sich teils stark an KEMPELENS Vorgaben orientieren, teils aber auch neue Wege gehen. Eine Zusammenschau der derzeit bekannten Sprachmaschinen-Repliken findet sich bei BRACKHANE (2011).

Glaubt man einem Hinweis aus »*Gehler's physikalischem Wörterbuch*« von 1825, so konnten bereits die zu diesem Zeitpunkt bekannten Repliken nicht mehr vollkommen überzeugen:

<sup>53</sup>Der Beschreibung bei CHAPUIS & GÉLIS zufolge war MÄLZELS »sprechender Kopf« allerdings gerade einmal zu »Mama« fähig [Chapuis & Gélis 1928, 210 f.].

»Manche nach seiner Anweisung gemachte Maschinen leisten nicht viel, und bringen kaum vernehmbare, der menschlichen Stimme nicht sehr ähnliche Töne hervor. Indess muss diejenige, welche der Erfinder selbst besass, verzüglich gewesen seyn, indem Collinson versichert, dass sie das von ihm aufgebene Wort: exploitation mit französischem Tone vernehmlich gesprochen habe.« [Gehler 1825, 655]

Über sämtliche historische Repliken ist heute nur noch wenig bekannt, bei allen reißt die Überlieferung recht rasch ab. Mit am besten dokumentiert ist die recht frühe Replik CHARLES WHEATSTONES (vgl. Kap. 4.3.2).

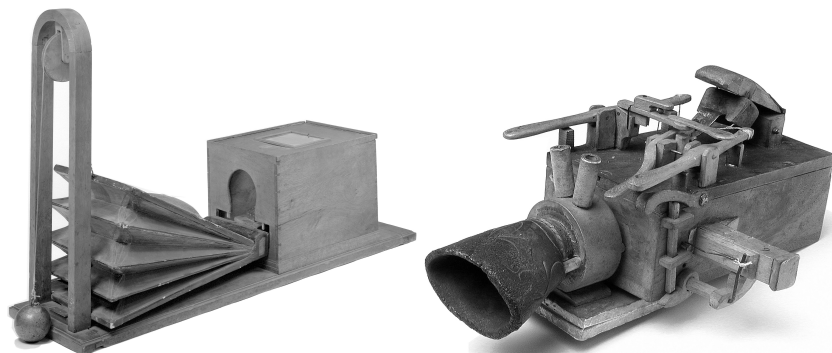
Vollkommen rätselhaft ist die Provenienz des im Deutschen Museum München (DMM) aufbewahrten Exemplars. Von zahlreichen Autoren (z. B. POMPINO-MARSCHALL 2005, 153; FELDERER 2005, 37) wird es als *das* Original KEMPELENS bezeichnet, was aus mehreren Gründen zumindest nicht in vollem Umfang zutreffend sein kann (vgl. hierzu [Brackhane 2011]). Es scheint sich bei diesem »Münchener Exemplar« in seinem heutigen Zustand vielmehr um ein Kompositum unterschiedlich alter Einzelteile zu handeln, dessen Kernbestand jedoch in Material und Faktur stark an die Sprachmaschine CHARLES WHEATSTONES erinnert (vgl. Abb. 3.7 sowie Abb. 4.4 und 4.5 S. 88/89.).<sup>54</sup> Die fehlerhafte Zuschreibung des Münchener Exemplars reicht zumindest bis zu CHAPUIS & GÉLIS zurück [Chapuis & Gélis 1928, 208].

Das DMM bemühte sich bereits zu Beginn seiner Aufbauphase 1905 in Person seines geistigen Vaters OSKAR VON MILLER (1855–1943) um die »Reste der Kempelen'schen Sprechmaschine« [Miller 1905], die in der Sammlung des physikalischen Instituts der Universität Wien vermutet wurden. Der Adressat des Briefes, PROF. DR. ANTON LAMPA (1868–1938), replizierte jedoch, die Maschine befinde sich

»in der Sammlung des physikalischen Institutes der K. K. technischen Hochschule in Wien (Prof. Dr. Gustav Jäger). In den Universitätssammlungen ist kein solcher Apparat vorhanden.« [Lampa 1905]

Woher das DMM schließlich sein Exponat erhielt, ist – wohl infolge der gravierenden Kriegsverluste – aus dessen Unterlagen heute nicht mehr ersichtlich. Die Recherchen des Autors in Wien ergaben jedoch, dass die Auskünfte LAMPAS offenbar nicht vollends zutreffend waren, da beispielsweise

<sup>54</sup>Umfangreiche Untersuchungen durch die Werkstätten des DMM, die dem Autor erst nach Fetigstellung des Manuskripts bekannt wurden, legen nahe, dass weite Teile der Maschine doch womöglich auf KEMPELEN selbst zurückgehen könnten.



**Abbildung 3.7:** Das »Münchener Exemplar« der KEMPELEN-Sprachmaschine.

die von ihm benannte Kontaktperson zumindest de jure zu diesem Zeitpunkt noch keinerlei Verfügungsrecht über die physikalischen Sammlungen der K. K. technischen Hochschule (nachmals TU) Wien hatte. Nach Auskunft von Frau Dr. JULIANE MIKOLETZKY, TU Wien, sind dort jedoch seit 1943 keine Unterlagen zu Bestandsbereinigungen der physikalischen Sammlung mehr vorhanden, so dass diese Spur auch dort nicht weiterverfolgt werden kann. Anscheinend wurde aber 1906 eine Sprachmaschine vom Wiener Konservatorium an das DMM abgegeben (pers. Mitteilung durch JULIANE MIKOLETZKY vom 19.02.2008, vgl. auch Reininger (2007), 330). Recherchen des Autors dort blieben jedoch wiederum ohne Erfolg.

Die folgende Tabelle 3.3 zeigt eine synoptische Übersicht über alle bislang bekannten Nachbauten der Sprachmaschine KEMPELENS. Die Bandbreite der Bauvarianten erstreckt sich hierbei von möglichst vorbildgetreuen Repliken bis hin zur künstlerisch-freien Adaption. Die meisten Nachbauten lassen sich unstreitig der ersten Kategorie zuordnen, während die Arbeit von FELDERER & STROUHAL das Paradebeispiel eines zugleich freien und sehr engen Umgangs mit der Thematik darstellt.

Für eine nähere Beschreibung und Abbildungen des »Münchner Exemplars« sowie der meisten anderen bekannten Repliken siehe [Brackhane 2011, 100 ff.].



**Tabelle 3.3:** *Bislang bekannte Sprachmaschinen-Nachbauten.*

Jahr	Ort	Erbauer	Vorbildgetreu	Zungen	Quelle
(vor) 1797	Jena	LODER	vermutlich	?	[Goethe 1893, 154]
1835	London	WHEAT- STONE	weitgehend	?	[Wheatstone 1879, 364]
vor 1906	Wien/ München	?	weitgehend	Elfenbein	pers. Mitt. S. BERDUX
1967	Utrecht	VAN DEN BROEKE	ja	Messing	[van den Broeke 1983]
1968	Paris	LIÉNARD	ja	Elfenbein	[Liénard 1969]
1993	York	HOWARD	teilweise	Messing	pers. Mitt. D. HOWARD
2001	Montluçon	DURIN	weitgehend	Elfenbein	pers. Mitt. J. JELTSCH
2001/ 2002	Budapest	NIKLÉCZY & OLASZY	ja	Elfenbein	[Nikleczy & Olaszy 2003]
2004	Wien	FELDERER & STROUHAL	nein	Messing	[Felderer & Strouhal 2004]
2007/ 2008	Saar- brücken	BAUER & BRACKHA- NE	ja	Elfenbein	[Brackhane & Trouvain 2008]
2008	Montluçon	JELTSCH	weitgehend	Elfenbein	pers. Mitt. J. JELTSCH
2009	Dresden	BRACKHANE & MAYER	ja	Elfenbein	BRACKHANE
2009	Paderborn	BRACKHANE & MAYER	ja	Elfenbein	BRACKHANE

### 3.5 Zusammenfassung

Noch bis weit ins 18. Jahrhundert hinein war die Szene der Sprache generierenden Maschinen von mehr oder weniger raffinierten Täuschungen bestimmt. Augenscheinlich gab es zeitweise beinahe eine regelrechte Inflation von »sprechenden Köpfen« und ähnlichen, oft mystifizierend inszenierten Attraktionen, die sich jedoch allesamt bald als unseriöse Konstrukte herausstellten. Entsprechend misstrauisch, ja offen ablehnend verhielt sich das Publikum denn auch, als insbesondere WOLFGANG VON KEMPELENS Sprachmaschine, von zahlreichen Publikationen wohl gut instruierter Autoren angekündigt, ins Licht der Öffentlichkeit trat. Mit KEMPELENS Billigung wurde zeitgleich der »mechanische Schachspieler« als hervorragend gemachter Taschenspielertrick »entlarvt«. Für die – wie wir heute wissen – seriöse Sprachmaschine war dies wahrlich keine Hilfe. Weder im deutschsprachigen Raum noch in Frankreich oder England konnte KEMPELEN auf seiner »Tournée« 1783-84 mit seiner Sprachmaschine die Anerkennung erhalten, die sie zweifelsohne verdient gehabt hätte. Nach etwa 1785 wird es publizistisch recht still um KEMPELEN, so dass wir nicht wissen, welche Wirkung die Veröffentlichung des »*Mechanismus*« 1791 auf die seinerzeit zahlreichen kritischen Stimmen hatte. Die Funktionsweise der Sprachmaschine war nun jedenfalls noch detaillierter offengelegt worden als es seinerzeit im Hinblick auf den »Schachtürken« der Fall war, womit KEMPELEN sicher auch die Betrugsverdächtigungen zu entkräften versuchte.

Die zweite weitreichende Entwicklung dieses Jahrhunderts, die »Vokalorgel« KRATZENSTEINS, fand bei weitem kein so lebhaftes Echo in der zeitgenössischen wissenschaftlichen Berichterstattung wie die Sprachmaschine KEMPELENS. Über die Gründe hierfür kann nur spekuliert werden. Zwar wurde KRATZENSTEINS Abhandlung über die Theorie der Vokalbildung bereits 1782 in einem französischen Medium veröffentlicht, doch sind seine Einlassungen über die Konstruktion der »Vokalorgel« außerordentlich dürftig, wie auch KRATZENSTEINS Vokaltheorie bereits bei seinen Zeitgenossen nicht unumstritten war. Da auch die offenbar umfangreichen Arbeiten des ABBÉ MICAL nur sehr begrenzten publizistischen Widerhall fanden, kann angenommen werden, dass es KEMPELEN schlicht besser als seine »Kollegen« verstanden hatte, die Mittel der Medien für seine Zwecke zu nutzen.

Ähnlich eigenartig unscharf wie bei den »Sprechenden Köpfen« MICALS bleibt auch das Bild von den Arbeiten DARWINS. Von beiden Apparaturen existieren nur wenige und detailarme zeitgenössische Beschreibungen. Beide sind verschollen, ein Schicksal, das sie mit den praktischen Arbeiten KRATZENSTEINS und KEMPELENS teilen. Während sich über die Beschaf-

fenheit der Erfindung MICALS sehr weitscheifig spekulieren lässt, legen die überlieferten Zeugnisse für DARWINS Maschine trotz ihrer Kargheit den Schluss nahe, dass es sich um eine recht unmittelbare Vorläuferkonstruktion der Sprachmaschine KEMPELENS gehandelt haben könnte.



# Kapitel 4

## 19. Jahrhundert: Epigonen & Erben

### 4.1 Letzte Pseudo-Sprachmaschinen

#### 4.1.1 L. Müller

Bereits in den Jahrzehnten nach dem bedeutungsträchtigen Jahr 1783 wurden Versuche unternommen, KEMPELENS »Auftrag« der Weiterentwicklung seiner Sprachmaschine nachzukommen [Kempelen 1791a, 456]. Gleichzeitig machten aber auch noch letzte Pseudo-Sprachmaschinen von sich reden. So trat 1788 ein Dr. L. MÜLLER (nähere biografische Daten unbekannt) u.a. in Erlangen mit einer Sprachmaschine an die Öffentlichkeit, deren in der »*Erlangischen Realzeitung*« beschriebene Vorführung stark an das von KEMPELENS »Schachtürken« überlieferte Procedere erinnert:

»Schon bey dem ersten Anblicke, wenn Herr M. den Kasten der Maschine (der etwa 2  $\frac{1}{2}$  Schuh breit und hoch, und 5 Schuh lang<sup>55</sup> seyn möchte) öffnet, und die Schubladen, worin die Sprachwerkzeuge, die Windladen, Hebel, Walzen, Federn, Register u. dergl. angebracht sind, herauszieht, um sie den Zuschauern vorzuzeigen, muß aller Verdacht einer Illusion oder Täuschung schwinden. Diese Schubladen füllen fast [!] den ganzen Raum des Kastens aus. Der Kasten selbst hat keinen bestimmten Stand im Zimmer, kann ohnbeschadet seiner Wür-

---

<sup>55</sup>Das entspricht ca. 76 bzw. 152 cm.

kung weggerückt, und nach Verlangen in eine andere Behausung gebracht werden. Zu beiden Seiten desselben stehen zwei Figuren in Lebensgröße, eine männliche und eine weibliche; beyde stützen sich mit dem einen Arm auf den Kasten und haben mit ihm keine andere Verbindung, als daß von dem inneren Mechanismus desselben durch beyde Arme Röhren bis an den Mund der Figuren fortgeführt sind. Es hängt von der Wahl der Zuschauer ab, welche von beyden Figuren sprechen soll. Alle Sprachwerkzeuge liegen im Kasten selbst, und die Figuren stehen nur pro Forma da, man kann sie wegnehmen, ohne daß die Wirkung unterbrochen wird. Vielmehr ist die Wirkung noch besser, wenn die Figuren weggerückt werden, weil die Töne alsdann unmittelbar aus der Maschine kommen. Man erstaunt über die Deutlichkeit der hervorgebrachten artikulierten Töne, wiewohl sich niemand überreden wird, daß die Menschenstimme vollkommen erreicht sey. Die Töne scheinen durch Flötenwerk hervorgebracht zu werden, welches man sich mit den Walzen vermittelst der Hebel in Verbindung gebracht denken muß. Aber ehe ein jeder Ton aus der Flöte kommt, muß er durch besondere Vorrichtungen, worinn das vorzügliche Geheimnis der Maschine besteht, modificirt und artikulirt werden. Die Arien, die die Maschine singt, scheinen ein für allemahl auf den Walzen angeordnet zu seyn, wie bey Spieluhren.« [Mayer 1788, 479]

In einer ebenfalls 1788 erschienenen Streitschrift bemühte sich deren anonymen Autor – ohne die MÜLLER'SCHE Konstruktion je zu Gesicht bekommen zu haben, allein aufgrund der ihm bekannten Berichte und logischer Beweisführung – diese als weiteren Betrugsfall zu entlarven [Anonymus 1788]. Einem Bericht der »*Berlinischen Monatszeitung*« aus dem Jahr 1789 zufolge behauptete MÜLLER selbst, dass seine Maschine namens Karoline durch Magnetismus mit Intelligenz und also auch mit Sprache ausgestattet sei [Karsten 1789, 182 ff.]. Verwunderlich ist die hier beschriebene, im Vergleich zum vorangehenden Zitat offenbar vollkommen andere äußere Gestalt der Vorrichtung, was die Vermutung zulässt, es habe sich um zwei unterschiedliche Konstruktionen gehandelt.

»Man sieht das Kopf- und Bruststück eines kleinen Frauenzimmers, welches ungefähr wie eine Zofe geputzt ist, und eine große Trompete vor ihrem Munde hält. Rechnet man den Kopf ab, so bleibt für die Höhe des Halses und untern Theiles der Puppe kaum mehr als 1 Fuß [ca. 30 cm], welches eben so für

die größte Breite gilt. Dieser solide Theil scheint, soweit ich ihn untersuchen konnte, von Leder, und so weit ausgestopft zu sein, als es der Raum zuläßt, welchen zwei kleine Blasebälge, die darin auf und nieder spielen, einnehmen, wozu unten ein bemerkbares Ventil angebracht ist. Diese Puppe ist durch einen mit Leder überzogenen krummgebogenen Haken, 4 Fuß über der Erde ungefähr, an einer Wand oder einem Pfeiler befestiget, und, wie es scheint, ganz frei. – Bei dieser Einrichtung nun, kann jeder aus der Gesellschaft ihr Fragen vorlegen, und sie beantwortet selbige ziemlich schnell, nicht undeutlich, und mit hellem Tone; auch bemerkt man keinen weitem Unterschied, wenn der Künstler selbst diese Fragen aufwirft. [...] Zuletzt muß Karoline noch ein Liedchen singen, welches denn freilich nicht ganz bezaubernd ausfällt. [...] Soviel ist nun hier gewiß, daß in der Puppe selbst kein menschliches Geschöpf Platz haben würde. Auch war es nicht möglich, daß bei der in Halle gewählten Einrichtung sich gleich hinter der Puppe ein Mensch befunden hätte, da sie an einer Spiegelwand zwischen zwei auf die Straße hinausgehenden Fenstern befestiget war. Dies brachte mich Anfangs auf die Vermuthung, daß Bauchsprache vielleicht das ganze Wunder bewirke. [...] Was mich indeß bald nöthigte, diese Vermutung ganz aufzugeben, war die Versicherung mehrerer glaubwürdiger Personen, daß Hr. Müller sogar das Zimmer verlasse, die Thür hinter sich zumache, und die Puppe demungeachtet auf vorgelegte Fragen wie vorhin antworte.« [Karsten 1789, 182 ff.]

Ein Bericht aus dem Jahre 1794 belegt, dass es sich um eine weitere Pseudo-Sprachmaschine handelte:

»Eines gewissen D. Müllers redende Maschine, womit derselbe vor einigen Jahren das Publicum öffte, ist in Franken glücklich entlarvt worden, obgleich der Eigenthümer das Zeugnis eines sonst sehr geschickten Mathematikers vor sich hatte, daß die ganze Sache auf künstlicher Mechanik beruhe. Die Maschine, mit welcher der Betrug gespielt wurde, steht noch in Nürnberg. Wer von der Kempelischen Maschine und der dabey gebrauchten Täuschung Nachricht geben kann, wird das Publicum verbinden, wenn er dies im Reichs-Anz. thut.« [Anonymus 1794, 703]

Offenbar hatte der Autor keine Kenntnis von der Existenz oder zumindest dem Inhalt des »*Mechanismus*«, so dass er bei KEMPELENS Konstruktion noch an einen Betrug glauben mochte.

#### 4.1.2 Heinrich Markus Brunner

Auch zehn Jahre darauf waren die Pseudo-Sprachmaschinen noch nicht aus dem Feld geschlagen. Selten indes dürfte es vorgekommen sein, dass sich der Verfertiger eines solchen Konstrukts selbst zu Wort meldete und die Möglichkeit echter Sprachmaschinen abstritt. 1798 schrieb der »Mechanikus« HEINRICH MARKUS BRUNNER (1745–1818) in seiner »*Ausführlichen Beschreibung der Sprachmaschinen oder sprechenden Figuren*«:

»Von mir kann man sich ganz vollkommen überzeugen, daß ich die Kunst Sprachmaschinen und Sprachwerke zu erklären ganz gründlich erlernt habe und verstehe, weil ich schon große und kleine dergleichen Sprachfiguren zum Privat-Scherz und Vergnügen zu verschiedenen Preisen verfertigt habe, und noch verfertige. Aus dieser meiner Beschreibung wird man finden, daß ich diese vier Maschinen so deutlich erklärt habe, daß jeder Leser sich überzeugen wird, daß es unmöglich ist, eine Sprachmaschine zu verfertigen, und daß sie durchgehends in der Producirung, Täuschung und Betrug sind.« [Brunner 1798, 11 f.]

Seinen Zweifel an einer wirklichen Sprachmaschine begründet er so, dass der Eindruck aufkommt, er könne von KEMPELENS oder einer anderen »echten« Sprachmaschine zumindest vage gewusst haben:

»Was nützt es, wenn man mit Aufwand, Fleiß, Mühe und Kosten, ein Orgelwerk mit Pfeifen, Blasbalg und Wellen [mechanische Umlenkvorrichtung in der Orgel] verfertigt, worauf man höchstens fünf Buchstaben in einerley Ton vernehmen könnte, und diese wären: A C D I K. – H gegen A ist ungleich schwerer, denn dieser Buchstabe wird von der Brust abgestoßen. Lippen-Buchstaben sind: B E F G M; mit geschlossenen Lippen: O P T U V W; Buchstaben von den Lippen und Zunge sind: L M [sic!] Q R S X Y Z. – Außer den ersten fünf Buchstaben ist es unmöglich, die übrigen vermög eines künstlichen Mechanismus anzubringen.« [Brunner 1798, 39 f.]

BRUNNERS grundsätzliche Kritik scheint berechtigt, war doch das Lautinventar sowohl bei MICAL als auch bei DARWIN und KEMPELEN nachweis-



lich stark eingeschränkt. Seine »Phonetik« ist jedoch weder von artikulatorischem Wissen getrübt noch nimmt sie Rücksicht darauf, dass nicht für das gesamte Alphabet eine direkte Buchstabe-Laut-Beziehung besteht. Allerdings ignorierte selbst KEMPELEN diesen Umstand geflissentlich (vgl. [Kempelen 1791a, 178 ff.]).

## 4.2 In Kempelens Schatten

### 4.2.1 Carl Salomon Warmholz

Weniges, und dies nur aus dritter Hand, ist über die zwei Sprachmaschinen des offenbar vielseitig begabten CARL SALOMON WARMHOLZ (1778–1854) aus Eisleben bekannt, der 1815 mit einer neuen Harfenmechanik [Nauwerck 1815, 548] und 1817 mit einem Flöte blasenden Automaten bekannt wurde [Krug 1817, 1292]. BOIS-REYMOND, der früheste Chronist dieser Apparatur, schreibt:

»Im November 1853 enthielten Berliner Zeitungen die Anzeige, dass der kürzlich in Eisleben verstorbene Mechanikus Warmholz, nebst anderen, nunmehr verkäuflich zu überlassenden werthvollen Kunstgegenständen, eine die menschliche Sprache täuschend nachahmende Sprachmaschine hinterlassen habe. Ich schrieb sogleich der Wittwe [sic!], mir nähere Auskunft über diese Maschine erbittend. Im Namen derselben, erhielt ich von Herrn Losse, Kaufmann, a. d. Helfta bei Eisleben, 9. Jan. 1854, eine Antwort, von der ich das Wesentliche hier anführe:

»Was die Sprachmaschine betrifft, so scheint diese unabhängig von der Kempelen'schen und Faber'schen gebaut zu sein. Wie mir der verstorbene Warmholz mittheilte, haben zu demselben Personen, welche die Faber'sche Maschine gehört haben, geäußert, dass seine Sprachmaschine die Faber'sche weit überträfe. Ich selbst kenne nur die Warmholz'sche und habe dieselbe einige Male durch ihn gehört. Es sind zwei Exemplare da, das eine mit 17, das andere mit 11 Tasten. Die Aussprache ist ziemlich deutlich und wird durch die Nachbildung der Sprachwerkzeuge, als Lippen, Zunge, Gaumen u. s. w. hervorgebracht. Die Sprache ist zwar modulirt, hat aber doch nicht die vielen und feinen Nüancirungen [sic!], welche die menschliche Sprache besitzt, so dass sie immer noch geisterhaftes besitzt, wie ich es bezeichnen möchte. Der bittende, schmeichelnde Ton war vorzüglich gelungen. Das Schlimmste hierbei ist nur, dass

derjenige, welcher die Maschine sprechen lassen will, die einzelnen Laute und die ganze Handhabung erst herausfingeriren [sic!] muss, da der Verstorbene eine specielle Anweisung dazu nicht hinterlassen hat, was den Verkauf ungemein erschwert. <« [Bois-Reymond 1862, 133 f.]

Zum Ankauf kam es schließlich ebenso wenig wie zur Anfertigung einer Kopie, so dass sich auch das Schicksal dieser offenbar interessanten Sprachsynthese im Dunkel der Zeit verliert. Zu fragen ist indessen, warum WARMHOLZ, wenn seine Sprachsynthese denn tatsächlich so überzeugend war, hiermit nicht genauso an die Öffentlichkeit ging wie Jahrzehnte zuvor mit seinem Flöte spielenden Automaten. Eine Antwort hierauf vermag womöglich der anonyme Autor eines ebenso umfassenden wie späten Nachrufs zu geben:

»Die letzte Erfindung von Warmholz war eine Sprachmaschine, die die menschliche Sprache kräftig und deutlich wiedergab, ja sogar sang. Wohl 3 Jahre lang hat er mehrere Leute an diesem Werke beschäftigt gehabt und es sich Tausende kosten lassen. In die Oeffentlichkeit ist er mit seiner Erfindung nicht getreten, weil ein Anderer in Wien mit einer solchen Erfindung ihm zuvorgekommen war [Kempelen?]. Ein Freund von ihm, Bergmeister Schulz, welcher die Maschine in Wien gehört hatte, suchte ihn zu beruhigen und erklärte ihm, daß die Maschine in Wien ein Kind, die seinige aber ein Mann und viel vollkommener sei. Aber nichts fruchtete, weder die Mahnung eines Ministers, noch die Zusicherung, daß der König die Maschine kaufen würde, konnten ihn bewegen, dieselbe zur Aufstellung zu bringen. Der Mechanismus brauchte nur noch in die Figur eingesetzt zu werden; aber trotzdem er ein sehr einfacher gewesen sein soll, hat später Niemand einen Laut wieder herausgebracht.

Als Warmholz im Herbst 1854 im 76. Lebensjahr verstorben war, wurden sämtliche Sachen, [...] an einen gewissen Friedländer in Nordhausen billig verkauft, der dann alles nach Berlin [siehe vorangehendes Zitat!] geschafft hat. « [Anonymus 1900, 459]

Dieser Nachruf wirft einige Fragen auf. WARMHOLZ starb 50 Jahre nach KEMPELEN und seine Sprachmaschine scheint seine chronologisch letzte Erfindung gewesen zu sein. Selbst wenn man davon ausgeht, dass sich diese Entwicklung in den 1840er Jahren vollzogen hat, würde dies bedeuten, dass

zu diesem Zeitpunkt noch eine Version den KEMPELENSCHEN Sprachmaschine in Wien existierte und bekannt war. Die Beschreibung der Synthese KEMPELENS als Kinderstimme scheint diese Vermutung zu stützen.

Indessen muss man bei einem Nachruf, der 50 Jahre nach dem Tode des Betreffenden entstand, wohl davon ausgehen, dass dessen Autor sich nur auf mündliche Überlieferungen aus zweiter und dritter Hand stützen konnte.

### 4.2.2 Leonhard Posch

Auch über die Konstruktion des Berliner Bildhauers LEONHARD POSCH (1750–1831) ist nur Weniges bekannt, das sich zudem auch noch zu widersprechen scheint. NIEMANN zitiert einen Artikel des Jahres 1807 aus dem Stuttgarter »*Morgenblatt für gebildete Stände*«:<sup>56</sup>

»In einem 3 Fuß langen und einen Fuß hohen und breiten<sup>57</sup> Kästchen befindet sich die nur 3 Zoll hohe und breite, 5 Zoll lange und 1  $\frac{1}{2}$  Zoll dicke<sup>58</sup> Sprechmaschine. Diese wird durch einen angesteckten Blasebalg, den man mit dem Ellenbogen des linken Armes drückt, in Bewegung gesetzt, und durch Berührung der angebrachten Klaves [Tasten] die Stimme hervorgebracht, welche der Stimme eines Kindes gleicht. Sie prononciert alle Buchstaben des Alphabets und eine große Anzahl deutlicher Worte in jeder Sprache und leistet in dieser Hinsicht mehr als selbst die Kempel'sche (so!) Der Direktor Denon<sup>59</sup> hat sie bei seiner Anwesenheit in Berlin gekauft und nach Paris gesandt.<sup>60</sup>

Danach handelte es sich um eine anscheinend ziemlich genaue, vielleicht in Einzelheiten verbesserte Nachbildung der KEMPELENSCHEN Maschine. Nach einer späteren Notiz der »*Vossischen Zeitung*« [vgl. [Bois-Reymond 1843], auch hier entsteht

<sup>56</sup>In der von Niemann angegebenen Ausgabe Nr. 212 vom 4. September 1807 konnte ein entsprechender Bericht nicht nachgewiesen werden.

<sup>57</sup>Je nach zugrunde gelegtem regionalem Fußmaß entspricht das ca. 90 cm x 30 cm.

<sup>58</sup>1 Zoll  $\approx$  30,0 mm, also 90,0 mm x 150,0 mm x 45,0 mm. Wie indessen die einzelnen Maße auf die Proportionen der Maschine zu beziehen sind, bleibt unklar.

<sup>59</sup>DOMINIQUE VIVANT DENON (1747–1825): 1804–15 Generaldirektor aller französischen Museen und Sachverständiger bei der Beschlagnahmung von Kunstgegenständen in den durch die Truppen Napoleon Bonapartes besetzten europäischen Gebieten.

<sup>60</sup>In den Beschlagnahmungslisten findet sich keinerlei Hinweis auf eine solche Konstruktion [Savoy 2010, Anhang, 391]. Ein regelrechter Ankauf der Maschine muss angesichts des Kontextes als unwahrscheinlich angesehen werden. Allerdings sprechen CHA-PUIS & GÉLIS explizit auch von einem Verkauf der Maschine und einer Ausstellung in Paris [Chapuis & Gélis 1928, 209].

der Eindruck einer genauen Kempelen-Replik] leistete sie aber erheblich mehr als jene, sodaß sie 1828 auf Anraten W. von Humboldts für die Kgl. Kunstkammer angekauft wurde. Dort war sie 1844 noch vorhanden und stand jedem, der sie zu spielen versuchen wollte, zur Verfügung. Die Kunstkammer wurde später [1876] aufgelöst und ihre Bestände den verschiedenen Museen überwiesen, aber weder die Generalverwaltung noch das Hohenzollern-Museum konnten mir etwas über den Verbleib der Maschine mitteilen.« [Niemann 1920, 27 f.]

Auf dieselbe Quelle bezieht sich augenscheinlich auch die nahezu wortgleiche Beschreibung bei CHAPUIS & GÉLIS [Chapuis & Gélis 1928, 209 f.]. POSCH hat angeblich KEMPELENS Bekanntschaft gemacht [Gessinger 1994, 627]. Träfe dies zu, wäre es eine einfache Erklärung für die Ähnlichkeit seiner Sprachmaschine mit der KEMPELENS. Die von GESSINGER angegebene Quelle bietet indes keinen Hinweis auf eine solche Bekanntschaft [vgl. Chapuis & Gélis 1928, 209]. BOIS-REYMOND hingegen, der sich auch den Ankauf durch die Königliche Kunstkammer zugute hält, hat erste Nachrichten über diese Maschine erst bei ihrem scheinbaren Wiederauftauchen 1828. Nach Auflösung der Kunstkammer 1876 kam die Maschine wohl ins Neue Museum [Bois-Reymond 1862, 131; Niemann 1920, 27; Gessinger 1994, 627], wo sich seine Spur verliert.<sup>61</sup> Es bleibt unklar, von woher die Maschine 1828 angekauft wurde. Laut BOIS-REYMOND erfolgte der Ankauf direkt vom Konstrukteur; dann müsste es sich aber wohl um eine andere als die um 1807 nach Paris transferierte Maschine gehandelt haben.<sup>62</sup>

<sup>61</sup> Recherchen bei den Staatlichen Museen zu Berlin blieben ohne Ergebnis. Die Bestände der Kgl. Kunstkammer wurden 1876 aufgeteilt und jeweils thematisch verwandten Berliner Museen übergeben, worüber jedoch ebenso wenig (mehr) Unterlagen existieren wie über die Bestände der Kgl. Kunstkammer vor ihrer Auflösung.

<sup>62</sup> Sollte die Konstruktion POSCHS 1807 tatsächlich als Beutekunst nach Frankreich verbracht worden sein, würde dies nicht einer gewissen Ironie entbehren. Denn ausgerechnet POSCH fertigte um 1806 für den seinerzeit auch in Deutschland sehr populären DENON ein Medaillon mit dessen Portrait [Savoy 2010, 123].

## 4.3 Die geistigen Erben

### 4.3.1 Robert Willis

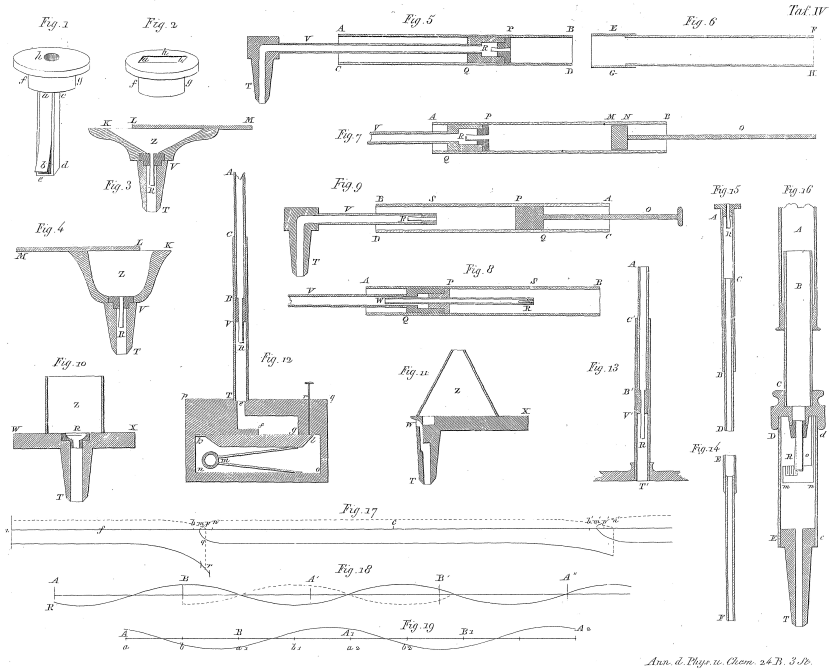
Wiewohl Priester, war der Cambridge-Professor ROBERT WILLIS (1800–1875) eine der prägenden Gestalten der englischen Ingenieurs- wie auch Sprachwissenschaften zu Beginn des 19. Jahrhunderts.<sup>63</sup> Bereits unmittelbar nach seiner Priesterweihe publizierte er in den Jahren 1829 und 1830 zwei Schriften zur menschlichen Sprache, darunter auch den uns hier besonders interessierenden Aufsatz »*On vowel sounds, and on reed-organ pipes*« (1830). WILLIS' Überlegungen waren offenbar so eindrucklich, dass bereits 1832 eine deutsche Übersetzung dieses Artikels erschien.

In seinen Überlegungen geht WILLIS davon aus, dass KEMPELENS Annahme, die Entstehung der Sprachlaute sei allein durch physiologische Gegebenheiten bedingt, falsch ist und wählt einen vollkommen anderen, experimentellen Ansatz:

»Ueberdies sind die Vocallaute eine Tongattung, die nachzuahmen nicht ganz außer dem Bereich der Kunst liegt, auch nicht ausschließlich den Organen des Menschen angehört. Musicalische Töne werden durch den Kehlkopf in der höchst möglichen Reinheit und Vollkommenheit hervorgebracht, und unsere besten Instrumente ahmen dieselben nur sehr unvollständig nach; allein, wem würde es wohl einfallen, in dem Kehlkopf eine Erklärung der Gesetze für die musicalischen Töne suchen zu wollen. Diese Betrachtungen veranlassten mich, bei Aufnahme dieser Untersuchung einen ganz anderen Weg [als KRATZENSTEIN und KEMPELEN] einzuschlagen, nämlich die Sprachwerkzeuge gänzlich zu vernachlässigen, und durch Versuche wo möglich zu bestimmen, welche Gestalten von Hohlkörpern, und welche Bedingungen zur Hervorbringung dieser Töne [Vokale] wesentlich seyen. Durch Vergleichung der erhaltenen Resultate mit den verschiedenen Lagen der Organe hoffte ich dann zu ermitteln, nicht nur, warum sie gerade diese oder jene Lage annehmen, sondern auch, welche Theile und Bewegungen we-

---

<sup>63</sup>Dass Geistliche der Anglikanischen Kirche zugleich Wissenschaftler, nicht selten von Rang, waren, ist ein typisches Phänomen des 18. und frühen 19. Jahrhunderts. Die Pfarrer der englischen Landgemeinden waren zu dieser Zeit meist umfassend gebildet und bezogen bei nicht allzu hoher Arbeitsbelastung ein Gehalt, das ihnen erlaubte, sich neben ihren Dienstpflichten intensiv ihren privaten Forschungen zu widmen [Strong 2007, 204 ff.].



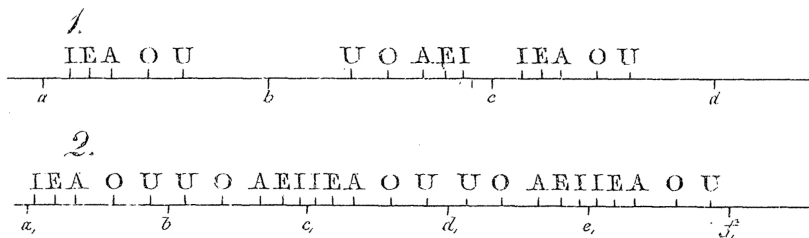
Ann. d. Phys. u. Chem. 24 B. 3. 18.

**Abbildung 4.1:** Bestandteile der Versuchsaufbauten von ROBERT WILLIS.  
[Willis 1832, Anhang].

sentlich für das Sprechen, und welche für andere Verrichtungen sind. « [Willis 1832, 399 f.]

Trotz dieses vollständig anderen Forschungsansatzes sah auch WILLIS in der Orgel und speziell in den *Zungenpfeifen* ein Analogon zur menschlichen Sprache bzw. deren Vokallauten: Die Blasebälge entsprächen der Lunge, die *Lingualpfeifen* den Stimmlippen und die *Resonatoren* dem Mundraum. Die unterschiedlichen Vokalraumformen ließen sich demnach durch verschiedengestaltige *Resonatoren* nachbilden, ein Gedanke, der stark von KRATZENSTEIN beeinflusst scheint [Willis 1832, 402].

In seiner umfangreichen Versuchsreihe griff er vor allem die Arbeiten KRATZENSTEINS auf und suchte sie experimentell fortzuführen. Mit seinen Ergebnissen gedachte er es den Orgelbauern zu ermöglichen, neue VH-Pfeifen zu bauen, die den Funktionsprinzipien der menschlichen Stimme entsprächen und somit auch tatsächlich vokalähnlich klingen sollten [Willis 1832, 161 f.]. Hierzu diente ihm eine Vielzahl verschiedener Ver-



**Abbildung 4.2:** »Vokalstrahle« zweier unterschiedlicher Konfigurationen der Vokalsynthese von Robert Willis [Willis 1832, Anhang].

suchsanordnungen, in denen er vorwiegend *durchschlagende Lingualpfeifen* mit unterschiedlichen *Resonatoren* als »Ansatzrohre« kombinierte (Abb. 4.1). Zunächst verwendete er hierzu analog zu KEMPELEN verschiedenförmige Trichter (Abb. 4.1, Fig. 3+4), die sich durch Schieber mehr oder weniger stark verschließen ließen. Mit diesem Aufbau wollte WILLIS bereits die Vokale u, o, a, e, i (in dieser Reihenfolge) erzeugt haben.<sup>64</sup> Weitere Versuche unternahm er mit verschiedenen zylindrischen Röhren, die in ihrer Länge teleskopartig veränderbar waren (Abb. 4.1, Fig. 5-16). Hierbei beobachtete er, dass bei sonst gleichbleibenden Parametern die Längenveränderung zur Wahrnehmung verschiedener Vokalqualitäten ausreiche. Mit sukzessiver Verlängerung des Rohres erhielt er die Vokalfolge i, e, a, o, u. Er stellte hierbei fest, dass sich ab einer bestimmten Länge die Vokalqualitäten wiederholten, allerdings in umgekehrter Reihenfolge (Abb. 4.2).

Für WILLIS war also nicht die physiologisch korrekte Nachbildung der anatomischen Gegebenheiten beim Menschen maßgeblich, wie es KEMPELEN und andere seinerzeit unternommen hatten, sondern ausschließlich ein überzeugendes Ergebnis der Synthese. Es handelte sich demnach um keine »indirekte Synthese«, bei der man durch die Nachbildung der relevanten menschlichen Organe hoffte, auch deren Funktion zu erhalten, sondern um eine »direkte Synthese«, bei der einzig das akustische Ergebnis von Relevanz war, nicht aber der Weg dorthin.

<sup>64</sup>Eine Ursache für Willis' Beobachtungen könnte das Phänomen F2' sein (vgl. [Stevens 1998, S. 289 f.]). Den initialen Hinweis hierauf verdanke ich Prof. Dr. JOHN OHALA (Berkeley).

### 4.3.2 Sir Charles Wheatstone

Nahezu zeitgleich mit WILLIS beschäftigte sich auch der britische Physiker SIR CHARLES WHEATSTONE (1802–1875, Erhebung in den Adelsstand 1868) – heute eher bekannt als Entdecker des dynamoelektrischen Prinzips neben WERNER VON SIEMENS (1816–1892) – mit mechanischer Sprachsynthese. In einem 1837 erschienenen Artikel befasste er sich geradezu in einem Rundumschlag sowohl mit dem »*Tentamen*« CHRISTIAN GOTTLIEB KRATZENSTEINS und dem »*Mechanismus*« WOLFGANG VON KEMPELENS als auch mit dem 1830 erschienenen, o. g. Artikel »*On the Vowel sounds...*« von ROBERT WILLIS.

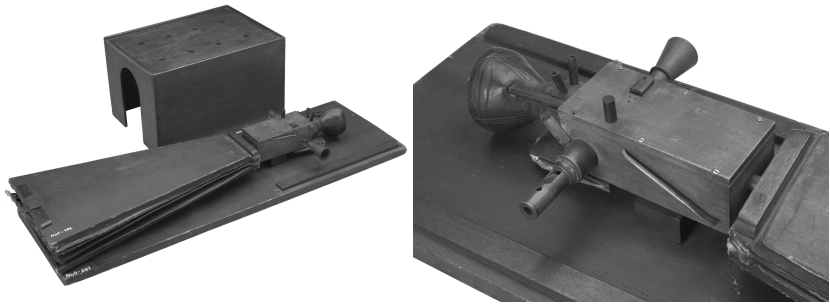
Hierbei setzte sich WHEATSTONE intensiv mit der Vorgeschichte der Entwicklungen KRATZENSTEINS und KEMPELENS sowie deren jeweiligen Konstruktionen auseinander. Einen Großteil des Artikels jedoch verwendet er darauf, die erst wenige Jahre zurückliegenden Untersuchungen WILLIS‘ detailliert zu schildern. Für unsere Zwecke ist allerdings eine (wohl von den Herausgebern angefügte) Fußnote von besonderem Interesse, die nahezu beiläufig erwähnt:

»A speaking machine, made by Professor Wheatstone from De Kempelen’s description, with some improvements, is at present in the collection of the Philosophical Instruments of King’s College, London.« [Wheatstone 1879, 364]

Laut DUDLEY & TARNOCZY wurde diese Replik indessen bereits 1835 auf einer Veranstaltung der »British Association for the Advancement of Sciences« in Dublin gezeigt [Dudley & Tarnoczy 1950, 164]. SIR RICHARD PAGET scheint der letzte Vertreter der Fachwelt gewesen zu sein, der die Maschine in Augenschein nehmen konnte und darüber berichtete [Paget 1930, 18]. Im Laufe der Recherchen zu seiner Magisterarbeit gelang es dem Autor, die Maschine, über deren Verbleib keine Klarheit mehr bestand, wieder ausfindig zu machen. Sie befindet sich heute – attribuiert als »Wheatstone’s artificial voice box« – in lädiertem Zustand in einem Magazin des Science Museum London. Im Rahmen dieser Dissertation konnten schließlich auch Fotografien dieser KEMPELEN-Replik angefertigt werden, von der bislang als einzige Abbildung die o.g. Skizze Pagets existierte. Allem Anschein fehlen das Balggewicht samt seiner Aufhängung sowie die Mechanik zur Erzeugung eines /r/-Lautes; im Übrigen macht die Maschine aber einen zwar leicht derangierten, jedoch vollständigen Eindruck (Abb. 4.3).<sup>65</sup>

<sup>65</sup>Sehr herzlicher Dank gebührt an dieser Stelle dem Institut für Phonetik an der Universität des Saarlandes in Person von Prof. Dr. BERND MÖBIUS, der sich großzü-





**Abbildung 4.3:** Die »Artificial voice-box« CHARLES WHEATSTONES.

Ein erster Blick zeigt zunächst in einigen Details eine verblüffende Ähnlichkeit zu dem im Deutschen Museum München vorhandenen Exemplar, was der seit langem schwelenden Diskussion über deren Provenienz neuen Schwung verleihen dürfte.<sup>66</sup>

Die Mundtrichter beider Repliken sind aus einem schwarzen, gummiartigen Material mit dekorativer floraler Prägung gefertigt. Ihre jeweilige Form ist aber merklich unterschiedlich, wobei die Londoner Version den bildlichen Vorgaben KEMPELENS sehr nahe kommt (vgl. [Kempelen 1791a, Tab. XXIII–XXV]).

Gerade diese Mundtrichter geben allerdings erhebliche Rätsel auf. KEMPELEN selbst beschreibt das zu verwendende Material als »ein Stück einer aus elastischem Gummi verfertigten gewöhnlichen Flasche, von der die untere Hälfte, und der engere Theil des Halses weggeschnitten ist« [Kempelen 1791a, 433]. Derartige Gummi- bzw. Kautschukflaschen (Abb. 4.4 rechts) waren indes zu dieser Zeit offenbar selten und wurden lediglich direkt in den kautschukanbauenden Regionen Amerikas hergestellt [Meyer 1863, 987]. Die Mundtrichter beider Repliken scheinen aus solchen Gummiflaschen hergestellt worden zu sein; insbesondere bei demjenigen

---

giger Weise bereitfand, die nicht unerheblichen Kosten für diese Bildbeschaffung zu übernehmen.

<sup>66</sup>So schreibt der anonyme Autor unter <http://de.wikipedia.org/wiki/Sprachsynthese>:

»CHARLES WHEATSTONE baute 1837 eine Speaking Machine, die auf diesem Entwurf [Kempelens eigener Beschreibung] basiert, ein Nachbau findet sich im deutschen Museum.« (zuletzt gesichtet am 27.06.2013)

Eine Formulierung, die leider nicht eindeutig ist. Nachfragen an den Autor blieben erfolglos. Über das Exemplar und dessen Herkunft im Deutschen Museum existieren mannigfaltige Behauptungen bis hin dazu, dass es sich um das Original Wolfgang von Kempelens handele, was aber als widerlegt gelten muss (vgl. [Brackhane 2011, 102]).

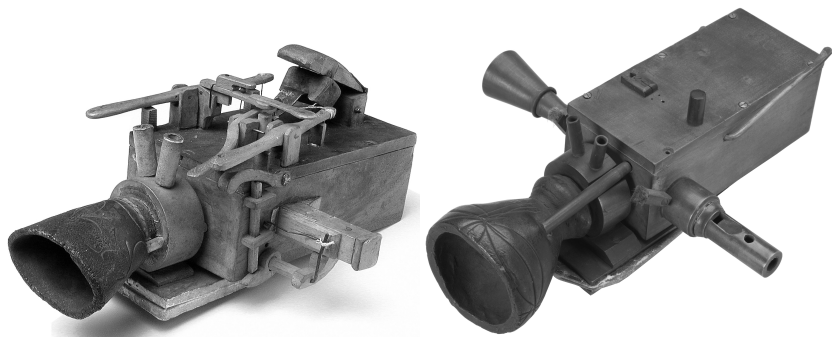


**Abbildung 4.4:** Vergleich der Sprachmaschinen-Mundtrichter aus München (links) und London (Mitte) mit einer Kautschuk-Flasche vom Amazonas (rechts, © Deutsches Museum München).

aus London sind deutliche Schnittkanten erkennbar. Beide Mundtrichter weisen eine florale Verzierung auf, die geprägt oder geritzt worden sein könnte. Angaben zu ggf. üblichen Verzierungspraktiken finden sich jedoch in der Fachliteratur nicht. Die Abbildung einer nahezu zeitgleich mit dem Mundtrichter in Cambridge entstandenen Kautschukflasche vom Amazonas legt jedoch eine Herkunft des Materials von dort nahe (vgl. Abb. 4.4 Mitte und rechts).

Holzart und Fertigung der Windladen weisen ebenfalls deutliche Übereinstimmungen auf. An zahlreichen Details wird jedoch bald deutlich, dass eine Abstammung aus ein und derselben Hand als ebenso unwahrscheinlich bezeichnet werden muss wie eine direkte Vorbild-Kopie-Beziehung (vgl. Abb. 4.5): Nicht nur die Formen der Mundtrichter sind unterschiedlich, sondern auch die Anlage und Konstruktion der Nasal-Öffnungen, des Bypass-Röhrchens, des Plosiv-Balgs, beider Frikativ-Generatoren sowie von deren Steuerung. In den meisten Aspekten erscheint die Londoner Maschine als deutlich vorbildgetreuer, einzig bei der Steuerung der Frikativ-Generatoren muss hier das Münchner Exemplar den Zuschlag erhalten.

Insgesamt macht besonders der direkte Vergleich mit der verlässlich auf 1835 datierbaren Wheatstone-Replik deutlich, dass es sich bei dem Münchner Exemplar ganz offenbar um eine recht späte Nachschöpfung – evtl. unter Verwendung älterer Teile – handeln muss, bei der mittels einer deutlich komplexeren Anlage und Steuerung versucht wurde, offenkundige Schwächen des KEMPELENSchen Originals zu beheben. Für eine moderne »Rekonstruktion« spricht auch, dass OSKAR VON MILLER in seinem Schreiben nach Wien explizit die »Reste der Kempelen'schen Sprechmaschine« thematisiert [Miller 1905]. Diese könnten nach ihrer Ankunft im DMM recht frei zu einem ansehnlichen Ganzen ergänzt und zusammen-



**Abbildung 4.5:** KEMPELEN-Replik aus München (links) und »Wheatstone's artificial voice box« (rechts).

gesetzt worden sein, wie es nach Auskunft von Frau DR. SILKE BERDUX (DMM) seinerzeit mit zahlreichen Exponaten des DMM geschah.

Worin die bei WHEATSTONE erwähnten Verbesserungen seiner Replik bestehen sollen, ist nicht vollends klar. Es könnte sich um eine Neukonzeption der Frikativgeneratoren handeln, die bei ihm einige Unterschiede zu KEMPELENS Beschreibung aufweisen. Im Kontrast mit denjenigen des Münchener Exemplars sind sie jedoch noch als sehr vorbildgetreu anzusprechen. Die zunächst unsinnig erscheinende Positionierung der Steuerungshebel für die Frikativ-Generatoren ist ganz offenbar auf eine fehlerhafte Ausrichtung, womöglich erst anlässlich der Fotografien, zurückzuführen: Diese müssten, statt nach hinten zu zeigen, auf den beiden gepolsterten Scharnieren zu liegen kommen, die beidseits der Windlade unmittelbar vor den jeweiligen Frikativ-Generatoren zu erkennen sind.

Eine kurze Beschreibung der Replik findet sich bei BRUCE, nach dem auch MELVILLE und ALEXANDER GRAHAM BELL äußerten:

»[We] heard it pronounce, in a very mechanical manner, a few simple words and sentences.« [Bruce 1973, 35]

### 4.3.3 Joseph Faber

Der 1843 erstmals bekannt gewordene Fall der »Euphonia« genannten Sprachmaschine eines gewissen JOSEPH FABER<sup>67</sup> muss als geradezu rätselhaft bezeichnet werden. Zwar fand diese in der Berichterstattung des 19.

<sup>67</sup>Sowohl die Lebensdaten als auch nähere Angaben zur Person sind unbekannt. Nach BRUCE starb FABER zwischen 1863 und 1870 [Bruce 1973, 82]. DUD-

Jahrhunderts lebhaft Aufnahme, ohne dass aber jemals konkrete Details über die Maschine oder ihren Schöpfer zu erfahren gewesen wären. Einzig wenigstens drei Abbildungen der Maschine existieren, von denen hier die zwei mutmaßlich detailliertesten wiedergegeben wurden (Abb. 4.6). FABER scheint WOLFGANG VON KEMPELENS Aufforderung am Ende des »*Mechanismus*« sehr ernst genommen zu haben. Dort nämlich ist zu lesen:

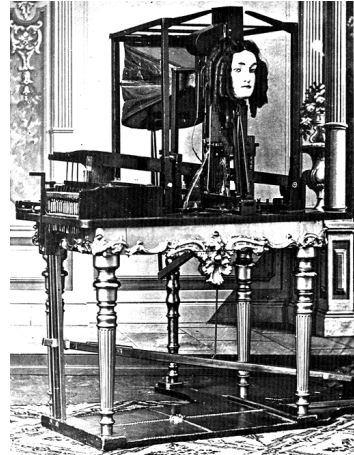
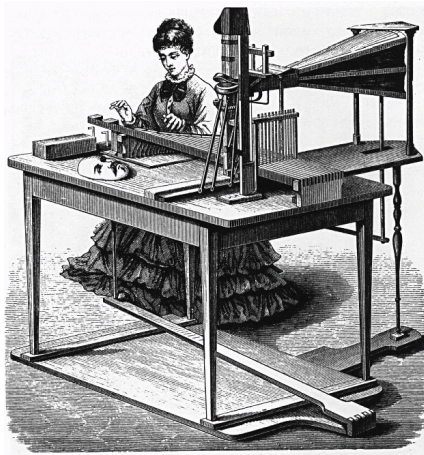
»Uibrigens [sic!] bin ich überzeugt, daß die Maschine ohne sonderliche Kunst mit Tasten, wie ein Clavier oder Orgel so einzurichten wäre, daß das Spielen auf derselben gegen der dermaligen Art [gegenüber der momentanen Art] jederman viel leichter fallen solle, aber das ist eben ein Schritt näher zur Vollkommenheit, den ich Einigen meiner Leser überlassen muß, die etwan dieser neuen, noch in ihrer Kindheit befindlichen Erfindung einige Aufmerksamkeit schenken, und sie durch ihr Nachdenken und Bemühen weiter fortrücken werden.« [Kempelen 1791a, 456]

Liest man vor diesem Hintergrund die zeitgenössischen Berichte über FABERS Erfindung, so sind diesen zumindest einige Anhaltspunkte zu entnehmen, die sowohl explizit wie auch implizit eine deutliche Beziehung zu KEMPELEN aufzeigen:

»Die Sprechmaschine, welche ein Künstler aus Wien, Faber mit Namen, hier kürzlich zu sehen oder vielmehr hören ließ, scheint uns in ihrer Art ein so gelungenes physikalisches Kunstwerk zu seyn, daß wir nicht umhin können, einige Worte darüber zu sagen, wiewohl leider nicht das Wesentliche ihrer Einrichtung berühren können. Die Maschine ist unstreitig ein bedeutendes vollkommener als die frühere von v. Kempelen, welche wir im Kings College durch Hrn. Wheatstone's Güte kennen lernten [Gemeint ist die durch CHARLES WHEATSTONE angefertigte Replik, siehe 4.3.2]. Während sich diese nur durch einen Trichter von Kautschuck vernehmen läßt, den man mit der Hand verschiedenartig verschließen und öffnen muß, besitzt die Faber'sche Maschine einen dem menschlichen nachgebildeten Mund mit Lippen und Zunge, aus demselben Material gefertigt, welcher bloß mit Hülfe eines Blasebalges und einer Klaviatur alle Buchstaben und Worte, folglich auch beliebige Sätze, in mehr als einer Sprache zwar nicht eben schön, aber

---

LEY & TARNOCZY bezeichnen ihn als »certain Professor Joseph Faber of Vienna« [Dudley & Tarnoczy 1950, 164].



**Abbildung 4.6:** FABERS »Euphonia« als Stahlstich (links) [Dudley & Tarnoczy 1950, 165] und als Fotografie (rechts) [Anonymus 2013a].

sehr verständlich hervorbringt. Ihre Stimme läßt sich sehr verstärken und schwächen, bis zum vollen Leise-Sprechen, auch vertiefen und erhöhen, daher denn auch ein Singen möglich ist. Die Klaviatur, welche 16 Tasten enthält, giebt die Vocale a, e, i, o, u, die Halbvokale r, l, w und die Consonanten f, s, s' (unser tsch) b, d, g. Die übrigen Consonanten werden aus diesen und mittels zweier Hülftasten hervorgebracht, von denen die eine die Stimmritze und die andere die Nase schließt oder öffnet. Die erste Hülftaste giebt die Aspiration unseres h, und verwandelt, gleich hinter g angegeben, dieses in k; die letztere, unmittelbar hinter b und d niedergedrückt, macht diese zu m und n. Immer muß den Consonanten ein Vocal angehängt oder vorgesetzt werden, wenn sie so zum Vorschein kommen sollen, wie wir sie gewöhnlich aussprechen. « [Poggendorff 1843, 175 f.]

Eine geradezu euphorische Beschreibung findet sich in der »*Vossischen Zeitung*« vom 24. Januar 1843. Der dort abgedruckte Text des preußischen Regierungsrats und Direktors der königlich-preußischen Kunstkammer FÉLIX-HENRI DE BOIS-REYMOND (1782–1864) legt zudem nahe, dass dieser entweder mit der Sprachmaschine POSCHS oder einer anderen Kempele-Replik recht vertraut gewesen sein muss:

»Unterzeichneter [...] erklärt hierdurch mit Vergnügen, daß er die von Herrn Faber aus Wien zu Stande gebrachte und von ihm in Berlin öffentlich producirt Sprechmaschine als eine höchst wichtige Erscheinung in diesem bis jetzt noch so sehr zurückgebliebenen Kunstfache erachtet. [...] Die Faber'sche Sprechmaschine übertrifft nun in jeder Beziehung bei Weitem die des Posch und die ihres berühmten ersten Erfinders v. Kempelen. 1) Die des von Kempelen ist ohne Vergleich schwerer zu spielen, weil der Spieler unmittelbar die einzelnen Theile der natürlich sehr zusammengesetzten Maschine in Bewegung setzen muß. Mit der flachen linken Hand spielt man die Vokale; mit den Fingern der rechten Hand, die Consonanten; zugleich muß der Ellbogen den Blasebalg treten [sic!]. Bei der Faberschen, wie bei dem Piano, brauchen hingegen nur Tasten angeschlagen zu werden, welche mit den von ihnen angegebenen einzelnen Buchstaben des Alphabets bezeichnet sind. 2) Die Consonanten sind bei der Faberschen Maschine, wenigstens für die deutsche Sprache, vollständig und übrigens viel reiner und bestimmter als bei der Kempelenschen, bei welcher einige ganz fehlen und andere mit keinem bestimmten Charakter herauszubringen sind. 3) Hinsichtlich der Vokale bleibt aber die Kempelensche Maschine vollends hinter der Faberschen. Sie wurden mittelst einer Art von (hölzenen) [sic!] Trichter hervorgebracht, dessen breite Oeffnung man mit der linken Hand nach Gutdünken mehr oder minder weit und nahe überwölbte. Wie Herr Faber die Vokale mit einem künstlichen Munde (von Kautschuk) hervorbringt, bleibt bis jetzt noch sein Geheimniß. Jedenfalls aber zeigt hierin seine Maschine vergleichsweise mit der v. Kempelenschen eine erstaunliche Vollkommenheit. Dies ist etwas ganz Neues, eine wahre Entdeckung im Gebiete der Mechanik und Akustik, worüber Herr v. Kempelen selbst Jahre lang vergeblich brütete. 4) Ist die Stimme der Faberschen Maschine zwar nicht die einer ausgebildeten Opernsängerin, so ist sie doch, wie bei der Kempelenschen, eine der menschlichen sehr ähnliche. Ihre Laute machen einen seltsamen Eindruck; man möchte beinahe sich ein Beben, eine mitfühlende Seele in der Maschine denken. Auch hier zeigt sich wieder eine entschiedene Ueberlegenheit der Faberschen Maschine über die Kempelensche. Diese ist eintönig, gesanglos wie eine kleine Kindertrompete. Nicht allein weiß die Fabersche die Wörter und

Sylben zu betonen, sondern sie erhebt sich sogar zum Gesang. [...]

In industrieller Hinsicht dürfte die Sprechmaschine manchem Künstler, Mechaniker und Uhrmacher neue Erwerbsquellen öffnen. Man denke sich eine stattliche Wand- oder Stutzuhr, welche die Stunden nicht schlägt sondern ausruft, einen Vers dazu singt oder nach Belieben bei Gelegenheiten ein feierliches Lebehoch: Mit Gott für König und Vaterland! erhebt. Wäre einmal die Fabrikation eingerichtet, so würde der Preis gar nicht so hoch zu stehen kommen. Jedenfalls aber wird die Sprechmaschine eine deutsche Erfindung bleiben, die abermals den deutschen Scharfsinn und Erfindungsgeist und zwar selbstredend [!] verkündigen wird. « [Bois-Reymond 1843]

Im zwei Jahre später erschienenen Ergänzungsband zur Neuauflage des »*Physikalischen Wörterbuchs*« von JOHANN S. T. GEHLER schließt eine knappe Zusammenfassung des POGGENDORFFSCHEN Artikels mit den Worten »Der eigentliche Mechanismus ist noch Geheimniß« [Gehler 1845, 580]. Man machte sich demnach Hoffnung, dass FABER auch dahingehend dem Vorbild KEMPELENS folgen würde, die Konstruktionsgrundlagen seiner Maschine früher oder später offenzulegen.

NIEMANN wartet – wiewohl erheblich später – mit einigen weiteren Details auf, die zum Teil geradezu verblüffende Parallelen zur Berichterstattung über KEMPELENS Sprachmaschine offenbaren:

»Wo ein vollkommenes Nachbilden nicht möglich war, suchte er [Faber] den Zweck durch ein mechanisches Ersatzmittel zu erreichen. Ein Blasebalg vertrat die Lunge und drängte durch ein Gummirohr die Luft in den künstlichen Kehlkopf, dessen Inneres nicht gezeigt wurde. Hier wurde durch Schwingungen dünner Lamellen der Laut auf ähnliche Weise hervorgebracht, wie in der Natur durch die Schwingungen der Kehlbänder. Die schwingende und dadurch tönende Luft strömt in die Mundhöhle, die durch nachgebildete Zunge [!], Lippen und Gaumen begrenzt und abgeteilt, von einem festen Oberkiefer und einem beweglichen Unterkiefer eingeschlossen und mit Gummi elasticum bekleidet ist. [...] Einige Wörter und Redensarten, auf die sich der Erfinder besonders eingeübt hatte, gelangen recht gut, andere machten anscheinend große Schwierigkeiten. [...]

Man sah übrigens in dieser Maschine durchaus nicht nur einen Gegenstand der Neugier, sondern erwartete von ihr mannigfachen Nutzen, praktischer und wissenschaftlicher Art. Für

die Ausbildung der Taubstummenlehrer, für die Sprachkunde, Phonetik, Akustik usw. [...] [vgl. das vorangehende Zitat aus der ›*Vossischen Zeitung*‹]

Faber führte seine Maschine Jahrzehnte hindurch bald hier, bald dort vor und verstand es, immer wieder die Aufmerksamkeit wenigstens vorübergehend auf den Apparat und das Problem der Sprechmaschinen überhaupt zu lenken. Zuletzt zeigte er sie anscheinend 1877 in Paris. [...] ›Für jemand, der unvorbereitet das Zimmer betritt, könnte die künstliche Stimme für die eines Taubstummen gelten, den man sprechen gelehrt hat.‹ Auch die Fabersche Maschine scheint verschollen zu sein. « [Niemann 1920, 28 ff.]

BOIS-REYMOND zufolge war FABER weder dazu zu bewegen, das Original seiner Maschine zu einem realistischen Preis von einer öffentlichen Sammlung ankaufen zu lassen, noch zu diesem Zweck eine Replik herzustellen [Bois-Reymond 1862, 131 ff.]. FABER führte seine Maschine auch in England vor. Der englische Impresario und Journalist JOHN HOLLINGSHEAD (1827–1904) beschreibt in seinen Memoiren eine Vorführung der Maschine geradezu schonungslos nüchtern:

»I paid my shilling and was shown into a large room, half filled with boxes and lumber, and badly lighted with lamps. In the centre was a box on a table, looking like a rough piano without legs and having two keyboards. This was surmounted by a half-length weird figure, rather bigger than a full-grown man, with an automaton head and face looking more mysteriously vacant than such faces look. Its mouth was large, and opened like the eyes of Gorgibuster in the pantomime, disclosing artificial gums, teeth, and all the organs of speech. There was no lecturer, no lecture, no music – none of the usual adjuncts of a show. The exhibitor, Professor Faber, was a sad-faced man, dressed in respectable well-worn clothes that were soiled by contact with tools, wood, and machinery. The room looked like a laboratory and workshop, which it was. The Professor was not too clean, and his hair and beard sadly wanted the attention of a barber. I have no doubt that he slept in the same room as his figure – his scientific Frankenstein monster – and I felt the secret influence of an idea that the two were destined to live and die together. The Professor, with a slight German accent, put his wonderful toy in motion. He explained its action: it was not necessary to prove the absence of deception. One key-



board, touched by the Professor, produced words which slowly and deliberately in a hoarse sepulchral voice came from the mouth of the figure, as if from the depths of a tomb. It wanted little imagination to make the very few visitors believe that the figure contained an imprisoned human – or half human – being, bound to speak slowly when tormented by the unseen power outside. No one thought for a moment that they were being fooled by a second edition of the "Invisible Girl" fraud. There were truth, laborious invention, and good faith, in every part of the melancholy room. As a crowning display, the head sang a sepulchral version of "God save the Queen," which suggested inevitably, God save the inventor. This extraordinary effect was achieved by the Professor working two keyboards – one for the words, and one for the music. Never probably, before or since, has the National Anthem been so sung. Sadder and wiser, I, and the few visitors, crept slowly from the place, leaving the professor with his one and only treasure – his child of infinite labour and unmeasurable sorrow. « [Hollingshead 1895, 67 ff.]

Anlässlich einer Vorführung der »Euphonia« in Philadelphia im Dezember 1845 untersuchte der amerikanische Physiker JOSEPH HENRY (1797–1878) die Maschine in der Erwartung, einen weiteren Betrugsfall vorzufinden. Er berichtet jedoch anschließend:

»I have seen the speaking figure [sic!] of Mr. Wheatstone of London but it cannot be compared with this which instead of uttering a few words is capable of speaking whole sentences composed of any words what ever. « [Hankins & Silverman 1995, 218]

Auch HENRY kommen ebenfalls Ideen zu konkreten Nutzungsmöglichkeiten der Maschine:

»I believe that many applications of Faber's machine could be imagined [sic!] in connection with the telegraph. The keys could be worked by means of electro-magnetic magnets and with a little contrivance not difficult to execute words might be spoken at one end of the telegraphic line which have their origin at the other. [...] Thus if an image of the kind were placed in the pulpit of several churches [...] the same sermon might be

delivered at the same minute to all.«<sup>68</sup> [Hankins & Silverman 1995, 218]

DUDLEY & TARNOCZY schließlich fassen zusammen:

»This machine is said to have been a big improvement over von Kempelen's, particularly from the standpoint of having a variable pitch that permitted singing.« [Dudley & Tarnoczy 1950, 164]

Nach FABERS Tod wurde die Maschine in den 1870ern durch einen angeheirateten Neffen weiterhin unter FABERS Namen präsentiert [Bruce 1973, 82]. So wenig Konkretes über die »Euphonia« auch bekannt ist, so viel unterschiedliche Abbildungen existieren von ihr, darunter offenbar sogar eine frühe Fotografie (Abb. 4.6 auf S. 91).

Es ist zu bedauern, dass sich scheinbar weder die »Euphonia« noch eine einigermaßen präzise Beschreibung erhalten haben. Zwar lassen die zeitgenössischen Berichte erahnen, dass auch diese Sprachmaschinen-Konstruktion im Endeffekt an sehr ähnlichen grundsätzlichen Problemen krankte wie es bereits diejenige KEMPELENS getan hatte, in deren Tradition sie auch steht. Doch scheint FABER hinsichtlich der Steuerung und der Stimmmodulation einige vielversprechende Ideen gehabt zu haben, die genauer zu kennen durchaus reizvoll wäre.

#### 4.3.4 Melville und Alexander Graham Bell

Nicht zuletzt die Rezeption der »Euphonia« FABERS und der KEMPELEN-Replik von CHARLES WHEATSTONE führte zumindest indirekt zu einer der nachhaltigsten Erfindungen des 19. Jahrhunderts: Nachdem ALEXANDER MELVILLE BELL (1819–1905), früher Phonetiker und Gehörlosenlehrer, bereits 1846 FABERS Konstruktion zu Gesicht bekommen und gehört hatte, konnten seine Söhne MELVILLE (†1870) und ALEXANDER GRAHAM BELL (1847–1922) die »Euphonia« 1863 in London besichtigen. Kurz darauf konnten sie zudem auch eigene praktische Erfahrungen mit der Sprachmaschinen-Replik CHARLES WHEATSTONES machen (siehe Kap. 4.3.2), zu dem ihr Vater ihnen den Kontakt vermittelt hatte. Nachdem WHEATSTONE den BELL-Brüdern daraufhin sein Exemplar von KEMPELENS »*Mechanismus*« zur Verfügung gestellt hatte, machten diese sich gemeinsam daran, auf Grundlage von KEMPELENS Erkenntnissen selbst eine Sprachmaschine zu entwickeln, die aber – um der menschlichen Anatomie

<sup>68</sup>Eine sehr ähnliche Idee für den Einsatz einer Sprachmaschine hatte bereits Jahrhunderte zuvor LEONHARD EULER (vgl. das Zitat auf S. 24).

noch näher zu kommen – beinahe wieder zu einem »sprechenden Kopf« geriet:

»Taking impressions from a human skull, Aleck made gutta-percha [chemisch dem Kautschuk verwandter Saft des Gutta-perchabaumes] replicas of the jaws, teeth, pharynx, and nasal cavities. His father guided him into substituting a purely functional resonance chamber for the nasal passages and restrained Aleck's theatrical yearning to give the machine a full cranium, a face, and even a wig. Aleck did provide it with soft rubber lips and rubber cheeks, which served a practical purpose as well as hiding the eerily disembodied grin of the gutta-percha teeth.

Aleck's most ingenious arrangements were the soft palate of rubber stuffed with cotton, worked by a lever from outside, and the tongue, sliced crosswise into half a dozen carefully shaped wooden sections, each raised or lowered separately from below, each padded with cotton, and all covered with a single sheet of rubber. Thus the position not merely of the tongue as a whole but also each key segment of it could be reproduced as in life.« [Bruce 1973, 36]

Die Brüder hatten also zumindest in der Theorie einen Weg gefunden, KEMPELENS halbherzigen Vorschlag für ein Zungenäquivalent so zu modifizieren, dass es dem Vorbild etwas näher kam (vgl. [Kempelen 1791a, 443 und Tab. XXVI]). Allerdings gerieten auch bei dieser Sprachmaschine die frühkindlichen Äußerungen à la »Mama!« am besten [Bruce 1973, 37].

Und wie um einer Tradition willen fehlt auch von dieser Sprachmaschine heute jede Spur.<sup>69</sup> Heute ist uns BELL, der sich unter dem Eindruck seiner »taubstummen« Mutter und der Tätigkeit seines Vaters zeitweilig selbst als Dozent für Sprechtechnik betätigte, selbstverständlich bekannt als Erfinder des Telefons. Wenig bekannt ist dafür, dass diese Erfindung in ihren Ursprüngen jedoch nicht als Kommunikationsmedium für jedermann ersonnen wurde, sondern als Hilfe für schwerhörige Menschen.

## 4.4 Zwischen Wissenschaft und Sensation

Trotz anderslautender Absichten ihres Schöpfers gelangte bereits die Sprachmaschine WOLFGANG VON KEMPELENS letztendlich im Beipro-

<sup>69</sup> Mehrfache Nachfragen des Autors bei den »Bell Laboratories« blieben ohne Antwort.

gramm der Vorführungen des »Schachtürken«, also quasi als Jahrmarkts-Attraktion, in die Berichterstattung und das kollektive Gedächtnis und nicht als wissenschaftlich-seriöse Studie zur Sprachproduktion. Auch spätere Sprachsynthese-Modelle gingen den Weg der Popularisierung. Zwei Aspekte dieses Graubereichs zwischen Wissenschaft und Sensation seien hier herausgegriffen.

#### 4.4.1 Johann Nepomuk Mälzel

Berühmt und der Nachwelt bekannt wurde der Pianist und Mechanikus MÄLZEL (1772–1838), Sohn eines Orgelbauers, durch die Mitentwicklung des Metronoms. Diese Arbeit an der Schnittstelle zwischen Musik und Mechanik war indessen typisch für ihn, auf den zahlreiche Musikautomaten wie das *Panharmonikon* (um 1800) oder der »mechanische Trompeter« (1807) zurückgehen. In ersterem verwendete MÄLZEL im Übrigen als einer der Ersten *durchschlagende Zungenpfeifen*, was in sofern nicht verwundert, als er nachweislich in engem Kontakt mit GEORG JOSEPH VOGLER stand, der sich zeitgleich mit MÄLZEL in Wien und Paris aufhielt [Kuffner 1808] (siehe auch 3.2.3).

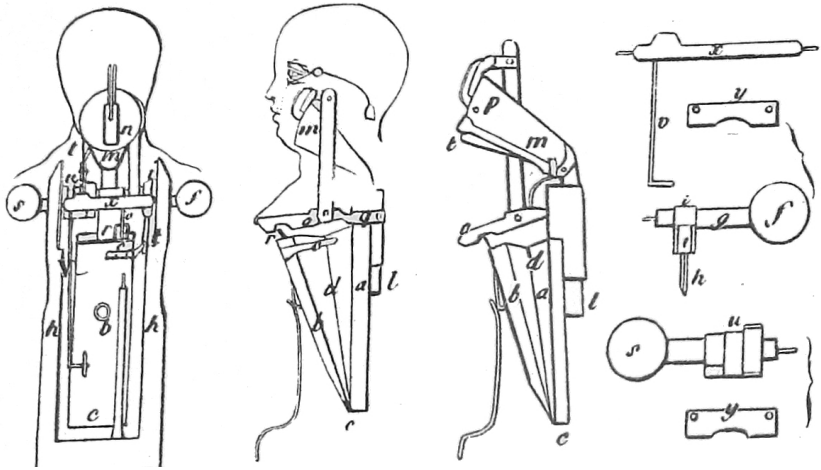
1806 erwarb MÄLZEL aus dem Nachlass WOLFGANG VON KEMPELENS den »Schachtürken« und begab sich mit diesem und weiteren mechanischen Kuriositäten auf eine Tournée ähnlich der, die KEMPELEN zwanzig Jahre zuvor unternommen hatte [Reininger 2007, 386]. Eine Bemerkung IGNAZ MOSCHELES' (1794-1870) vom Dezember 1823 in den Konversationsheften LUDWIG VAN BEETHOVENS (1770-1827) (für den MÄLZEL u. a. mehrere Hörrohre angefertigt hatte) lässt indessen vermuten, dass MÄLZEL auch über KEMPELENS Sprachmaschine zumindest recht genau Bescheid gewusst haben muss:

»Der arme Maelzel ist in Paris [...]. Jetzt macht er Puppen, die *Papa* und *Mama* aussprechen.« [Köhler & Herre 1968, 293]

MÄLZEL scheint offenbar KEMPELENS Mechanimus in eine praxistaugliche Puppenform transformiert zu haben, ein Vorsatz, den KEMPELEN selbst offenbar bereits in den 1780er Jahren gehabt hatte [Windisch 1783a, 50]. Ein entsprechendes Patent erhielt er 1824:

»It consisted a bellows, reed, and cup-shaped resonator that would be covered and opened mechanically, thus producing the words ›mama‹ and ›papa‹.« [Hankins & Silverman 1995, 213]

Eine ausführlichere Beschreibung samt Skizze (Abb. 4.7) bringen CHAPUIS & GÉLIS:



**Abbildung 4.7:** MÄLZELS »sprechende Puppe« [nach Chapuis & Gélis 1928, 210].

» Tout le monde connaît les poupées parlantes qui sont plutôt un jouet qu'une machine d'ordre scientifique. Plusieurs ouvrages donnent la description de la figurine de Maelzel pour laquelle celui-ci prit un brevet d'invention en 1824. Le principe du mécanisme est assez simple. Un soufflet envoie de l'air dans un tuyau à anche qui émet un son que l'on peut assimiler à la voyelle *a*; une soupape s'ouvrant lorsqu'on veut faire parler le jouet, produit une arrivée brusque de l'air qui rappelle *p*; la poupée >dit< ainsi *papa*. Un autre dispositif permet de réduire l'émission, ce qui donne à peu près *m*; la poupée >dit< alors *mama*. « [Chapuis & Gélis 1928, 210 f.]

» Alle Welt weiß, dass sprechende Puppen eher ein Spielzeug sind als eine wissenschaftliche Maschine. Mehrere Bücher geben die Beschreibung der Figur von Mälzel, für die er 1824 ein Patent erhielt. Das Prinzip des Mechanismus ist sehr einfach. Ein Blasebalg schickt Luft in eine Zungenpfeife, die einen Ton hervorbringt, der an den Vokal *a* erinnert; ein Ventil, das man öffnet, wenn man das Spielzeug sprechen lassen will, erzeugt einen abrupten Luftfluss, was an ein *p* erinnert; die Puppe >sagt< *papa*. Eine andere Einrichtung verringert den Luftfluss, was zu etwas wie *m* führt; die Puppe >sagt< dann *mama*. «

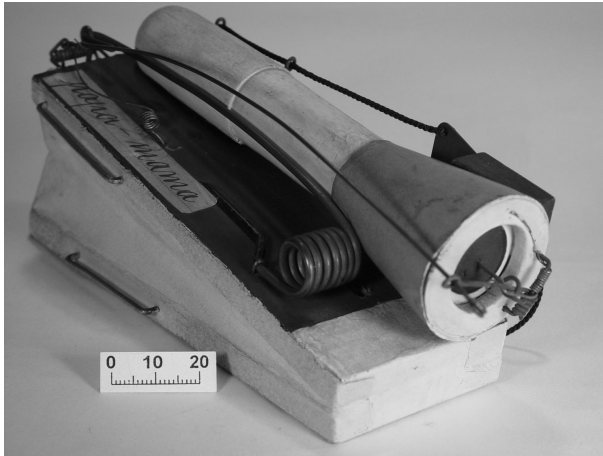
#### 4.4.2 Die thüringischen »Stimmenmacher«

Die Puppenmechaniken MÄLZELS waren der Beginn einer Entwicklung, die zum Ende des 19. Jahrhunderts in eine regelrechte Industrie münden sollte. Im sog. »Thüringer Spielzeugland«, in dem sich bereits seit dem 17. Jahrhundert ein Schwerpunkt der Spielzeugfertigung befand, wurden im Laufe des 19. Jahrhunderts auch Vorrichtungen für Puppen gefertigt, die diesen zumindest rudimentäre Sprachäußerungen erlaubten. Hierzu entwickelte sich sogar ein eigener Berufsstand, der des »Stimmenmachers«. Die so gefertigten Stimm-Mechaniken weisen konstruktiv ganz deutliche Parallelen zu den Vorarbeiten KEMPELENS und KRATZENSTEINS auf. Es handelt sich hierbei um einzelne hölzerne *Zungenpfeifen*, die liegend auf einem passenden kleinen Blasebalg montiert sind und jeweils einen Einzelvokal oder einen unkomplexen Zweisilbler wie »Mama, Papa, Emma, hurra« zu produzieren imstande sind [Hoffmann & Mehnert 2007] (Abb. 4.8).

Der Mediziner JOHANNES KESSEL (1839–1907) griff an der Schwelle zum 20. Jahrhundert schließlich eine Konstruktion HUGO HÖLBES auf, deren Entwicklung bis in die 1870er Jahre zurückreicht und knüpfte damit unmittelbar an das ehrgeizige Ziel KEMPELENS an: Mithilfe der von HÖLBE gebauten Stimm-Mechaniken sollte Schwerhörigen eine Unterstützung im Sprachunterricht gegeben werden. Zwar konnten die Stimm-Mechaniken unkomplexe CV-Silben synthetisieren, jedoch waren sie nicht in der Lag, die zur Artikulation dieser Silben notwendigen Bewegungen von Zunge, Lippen etc. zu visualisieren. Hinzu kam eine von KESSEL selbst eingeräumte mangelhafte Synthesequalität [Hoffmann & Mehnert 2007].

Dennoch stellte KESSEL die Entwicklung 1899 der Fachöffentlichkeit vor [Kessel 1900]. Er hatte erkannt, dass sich die seinerzeit pauschal als »Taubstumme« bezeichnete Gruppe von Kranken sehr heterogen darstellte. In ihr waren neben »Schwachsinnigen und Idioten« [Kessel 1900, 28] auch solche Kranke zu finden, deren einziges Handicap in Bezug auf das Erlernen von Sprache und den Erwerb von Wissen eine rein organische Taubheit oder gravierende Schwerhörigkeit war [Kessel 1900, 28]. Für die letzte Gruppe der hochgradig Schwerhörigen hatte KESSEL den Wunsch, diese unter Anwendung geeigneter Therapien soweit zu rehabilitieren, dass ihr Aufenthalt in den Taubstummenanstalten nicht weiter vonnöten und ihnen ein selbstbestimmtes Leben möglich sei. Da der Sprechunterricht bislang für das Lehrpersonal außerordentlich kräftezehrend sei, entwickelte er als Hilfsmittel für den Sprechunterricht die Stimm-Mechaniken.

Jede dieser Mechaniken bestand aus einem kleinen als *Keilbalg* konstruierten Blasebalg mit einer darauf montierten *Lingualpfeife*. Diese dürf-



**Abbildung 4.8:** Die Stimm-Mechanik für »Papa/Mama« [Hoffmann & Mehnert 2007].

te *durchschlagend* konstruiert worden sein. Synthetisiert werden konnte entweder einer der fünf »Grundvokale« oder /r/ oder Lautverbindungen wie »au« »Papa/Mama«, »Emma«, »Miau« und »Hurrah« (vgl. Abb. 4.8) [Hoffmann & Mehnert 2007].

## 4.5 Zusammenfassung

War das 18. Jahrhundert bestimmt durch das Auftreten mehrerer schillernder Forscherpersönlichkeiten, von deren Arbeiten insbesondere diejenigen des ABBÉ MICAL und WOLFGANG VON KEMPELEN ihren Niederschlag in zeitgenössischen Publikationen fanden, ist das 19. Jahrhundert geprägt von einer stark epigonalen Entwicklung. So ist sowohl das Auftauchen letzter Pseudo-Sprachsynthesen zu verzeichnen wie auch eine Vielzahl von Arbeiten – sowohl theoretischer wie praktischer Natur –, die die vorangegangene stürmische Entwicklung nachzuvollziehen und abzuschließen versuchen. Die ganz offenbar auf KEMPELENS Sprachmaschine beruhenden Arbeiten von POSCH, WALMHOLZ, FABER und BELL sind ein eindruckliches Zeugnis dafür, welch weite Verbreitung die Gedanken dieses Sprachsynthesepioniers trotz der mit 195 Exemplaren recht überschaubaren Auflage seines »*Mechanismus*« gefunden hatte.

Andererseits waren mit den Forschungen von WILLIS erste Ansätze geschaffen, die Arbeiten KRATZENSTEINS kritisch zu untersuchen und in

Kontrast zu den ebenfalls kritisch bewerteten Syntheseergebnissen KEMPELENS zu setzen. Das Verschwinden der »Euphonia« FABERS muss als überaus bedauerlich bewertet werden, scheint mit dieser Konstruktion doch tatsächlich eine deutliche Fortentwicklung der Arbeiten KEMPELENS gelungen zu sein, was auch die recht umfangreiche und nahezu durchweg positive Berichterstattung vermuten lässt.

Schließlich und endlich gelangt in diesem Zeitraum die Sprachsynthese – wenn auch in höchst rudimentärem Umfang – in die allgemeine Wahrnehmung, indem sie als Bestandteil von Spielzeug, insbesondere Puppen, Verwendung findet.



# Kapitel 5

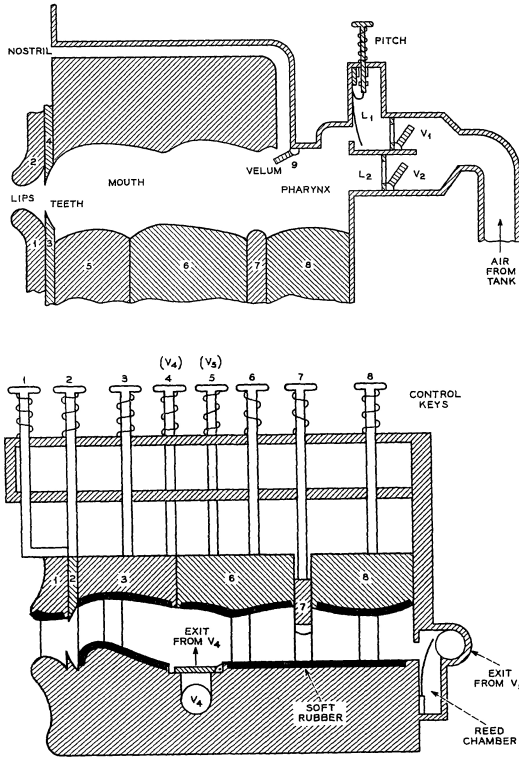
## 20. und 21. Jahrhundert: Letzte Innovationen

### 5.1 Sprachsynthesen aus den »Bell Laboratories«: Robert R. Riesz und Homer Dudley

Das 20. Jahrhundert ist gekennzeichnet durch sprachsynthetische Entwicklungen, die oft einen radikalen Schnitt mit den vorangegangenen Entwicklungen bedeuten und sich auf gänzlich neue Ansätze verlegten. Dies war natürlich nicht zuletzt möglich durch die enormen technologischen Fortschritte.

Die zu dieser Zeit entstandenen Sprachsynthesekonzepte lassen sich grob in zwei Gruppen einteilen. Da sind zunächst die weiterhin stark physiologisch geprägten Ansätze, die versuchten, den menschlichen Vokaltrakt so getreu wie möglich nachzubilden und auf diese Weise in einem zweiten Schritt durch einen artifiziellen Vokaltrakt auch artifizielle Sprache zu erhalten (»indirekte Synthese«). Auf der anderen Seite stehen die Konzepte, die sich der mittlerweile gezielt benutzbaren Elektrik und später Elektronik bedienen und deren Konzepte zur Spracherzeugung keinen physiologischen Hintergrund mehr haben. Aus beiden Lagern seien hier die bemerkenswertesten Exponenten, die beide in den »Bell Laboratories« entstanden, herausgegriffen.

Der amerikanische Physiker ROBERT R. RIESZ (1903–1974) präsentierte 1937 als Ergebnis seiner Arbeit bei den Bell Laboratories einen »Mechanical Talker« mit aus Gummi und Metall konstruiertem, in einzelnen Segmenten variierbarem Vokaltrakt, der auf Kempelens Entwicklung basierte.



**Abbildung 5.1:** Zwei Schemata des »Mechanical Talker« von RIESZ [Flanagan 1965, 208].

Eine gewisse Tradition ist auch hier gewahrt, da sich erneut Reminiszenzen zu einem Musikinstrument finden, wenn in diesem Fall auch nicht zur Orgel: Die Steuerung geschah über zehn Pumpventile ähnlich denen einer Trompete. Die verständlichste Äußerung soll das englische Wort »Cigarette« gewesen sein. Ebenso wie bei der Sprachmaschine KEMPELENS bedurfte es vonseiten des »Sprechers« einiger Übung, um die komplexe Steuerung so zu bedienen, dass eine dem Vorbild ähnliche Artikulationsfolge entstand [Flanagan 1965, 207 ff.].

FLANAGAN schreibt weiter:

»[zu Abb. 104oben:] Air under pressure is brought from a reservoir at the right. Two valves, V1 and V2 control the flow. Valve V1 admits air to a chamber L1 in which a reed is fi-

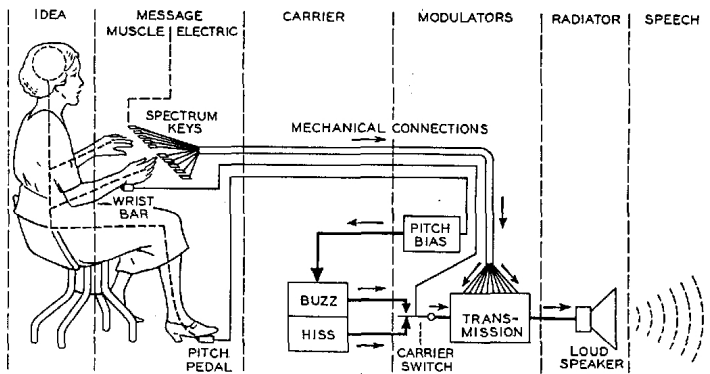


Fig. 8—Schematic circuit of the voder.

**Abbildung 5.2:** Funktionsschema des »Voder« [Lee 2013].

xed. The reed vibrates and interrupts the air flow much like vocal cords. A spring-loaded slider varies the effective length of the reed and changes its fundamental frequency. Unvoiced sounds are produced by admitting air through valve V2. The configuration of the vocal tract is varied by means of nine movable members representing the lips (1 and 2), teeth (3 and 4), tongue (5, 6 and 7), pharynx (8), and velar coupling (9).

To simplify the control, Riesz constructed the mechanical talker with finger keys to control the configuration, but with only one control each for lips and teeth (i.e., members 1-2 and 3-4 worked as pairs). The simplified arrangement with control keys is shown in the figure below [hier Abb. 5.1 unten]. The dark surface regions indicate soft rubber linings to accomplish realistic closures and dampings. Keys 4 and 5 operate excitation valves V4 and V5, arranged somewhat differently from V1 and V2 in the figure above [Abb. 5.1 oben]. Valve V4 admits air through a hole forward in the tract (below element 6) for producing unvoiced sounds. Valve V5 supplies air to the reed chamber for voiced excitation. In this case pitch is controlled by the amount of air passed by valve V5. When operated by a skilled person, the machine could be made to simulate connected speech. « [Flanagan 1965, 169 f.]

Bemerkenswerterweise wahrt RIESZ bei aller Innovation auch bei einem zentralen Aspekt die Tradition: Die Simulation der Stimmlippen geschieht nach wie vor über eine – offenbar durchschlagende – Zungenpfeife, wobei die hier präsentierten Schemata keinerlei sichere Aussagen über deren genaue Konstruktion zulassen. Auch hinsichtlich des Schicksals der Konstruktionen wurde allem Anschein nach der Tradition Genüge getan: Über ihren Verbleib gibt es keinerlei Informationen; mehrfache Nachfragen bei den »Bell Laboratories« über den Verbleib der RIEZ'schen Arbeiten blieben ohne Antwort.

Einen vollständig anderen Weg ging der Akustik-Ingenieur HOMER DUDLEY (1896–1987). Er präsentierte auf der New Yorker Weltausstellung von 1939 den »Voder« (»Voice operating demonstrator«), den er aus dem wenige Jahre zuvor ebenfalls in den »Bell Laboratories« entstandenen »Vocoder« (»Voice encoder«) weiterentwickelt hatte (Abb. 105). Diente der »Vocoder« noch der elektrischen und verschlüsselten Übermittlung von Sprache via Telefon, war es mit dem »Voder« erstmals möglich, Sprache auf elektrischem Wege zu synthetisieren. Um die Apparatur bedienen zu können, bedurfte es allerdings einer einjährigen Ausbildung.

Über einen Sinusgenerator wurde ein Ausgangssignal für stimmhafte Laute erzeugt, dessen Frequenz mittels eines Fußschalters modifiziert werden konnte. Für stimmlose Laute existierte ein Rauschgenerator. Die Umschaltung zwischen beiden geschah durch einen Handgelenkschalter. Durch zehn vom Benutzer über »spectrum keys« zu bedienende Bandpassfilter konnte das so generierte Signal beeinflusst werden, so dass ganze Sätze synthetisiert werden konnten [Lee 2013].

## 5.2 Vokalsynthesen mit Labialpfeifen

### 5.2.1 Dayton Clarence Miller

Hatten bereits zuvor vereinzelt Forscher versucht, auch mit *labialen* Orgelpfeifen eine Sprachsynthese zu bewerkstelligen [Miller 1905, 246], so sorgte die rasante technische Entwicklung zu Beginn des 20. Jahrhunderts für die Voraussetzungen, um diesen Gedanken tatsächlich in die Tat umzusetzen. Erst durch die Möglichkeit, einzelne Teiltöne eines komplexen Klanges grafisch darzustellen und zu messen, konnte die komplexe Struktur der Vokale näher erforscht werden. Als ein Resultat dieser Forschungen entwickelte DAYTON CLARENCE MILLER (1866–1941) eine auf *Labialpfeifen* basierende Vokalsynthese. Auf der Erkenntnis fußend, dass die Klanggestalt eines Vokals und seine Unterscheidbarkeit von anderen Vokalen eng mit

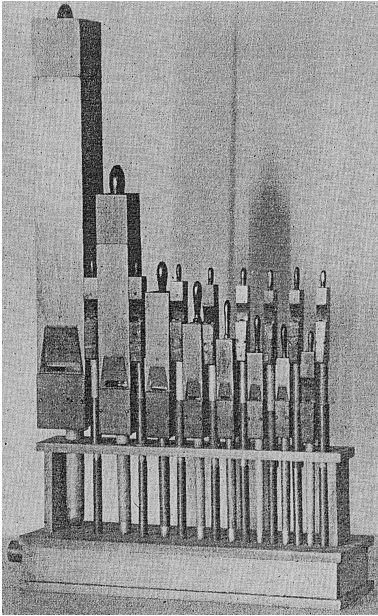


**Abbildung 5.3:** *Tibia clausa* (nach Anonymus 2012b).

der komplexen harmonischen Zusammensetzung seines Klanges verknüpft ist, stellte er für mehrere Vokale des amerikanischen Englisch Gruppen aus *Labialpfeifen* zusammen, die – kontrolliert durch den Vergleich mit Oszillogrammen »echter« Vokale – diese nachahmen sollten.

MILLER erkannte, dass zum Zwecke der Zusammensetzung eines komplexen Klanges aus mehreren Einzeltönen nur solche Tonerzeuger dienlich sein konnten, die neben ihrem Grundton möglichst keinerlei Obertöne erzeugen. Als hierfür geeignet stellte sich eine Bauform von *Labialpfeifen* heraus, die sich insbesondere in den amerikanischen Kinoorgeln findet und dort die Bezeichnung *Tibia clausa* trägt (Abb. 5.3). Es handelt sich um hölzerne, *gedeckte* Pfeifen mit verhältnismäßig großer Weitenmensur und hohem *Aufschnitt* von ca.  $\frac{1}{2}$ , die mit einem (verhältnismäßig) niedrigen Luftdruck angeblasen werden und in deren Klang laut MILLER der Grundton bis zu 99% des Gesamtspektrums ausmacht [Miller 1922, 247].

MILLER stellte nun – je nach Vokal – verschieden große Gruppen aus Pfeifen dieser Bauart zusammen. Hierbei war einerseits zu beachten, dass die Pfeifen so dicht als irgend möglich beieinander stehen und sich andererseits nicht gegenseitig in ihrer Absprache behindern. Hierzu wurde eine höhenversetzte Aufstellung gewählt (Abb. 5.4). Für /u/ waren sechs Pfeifen (also ein Klang bestehend aus sechs Partialtönen) notwendig, für /a/ zehn, für /æ/ bereits sechzehn. Die Einstimmung der Grundfrequenz erfolgte über eine Stimmgabel, die übrigen Teiltöne wurden bei ständiger optischer Kontrolle durch das *Phonodeik*, einen Vorläufer des Oszilloskops, beige stimmt und an die ebenfalls per *Phonodeik* aufgezeichneten Referenzvokale angepasst (Abb. 5.5). Je komplexer der synthetisierte Klang war (aus je mehr Einzelpfeifen er also bestand), desto schwerer war eine vollständig befriedigende Abstimmung aller Partialtöne aufeinander zu erreichen [Miller 1922, 249]. Diese Imperfektion in der Abstimmung der Pfeifen untereinander schlägt sich in einer sukzessiven Veränderung des *Phonodeik*-Signals über die Zeit und damit einer Entfernung vom menschlichen Vorbild nieder (vgl. Abb. 5.5). Dennoch resümiert MILLER:



**Abbildung 5.4:** MILLERS Set von sechzehn Labialpfeifen für den Vokal /æ/ [Miller 1922, 248].

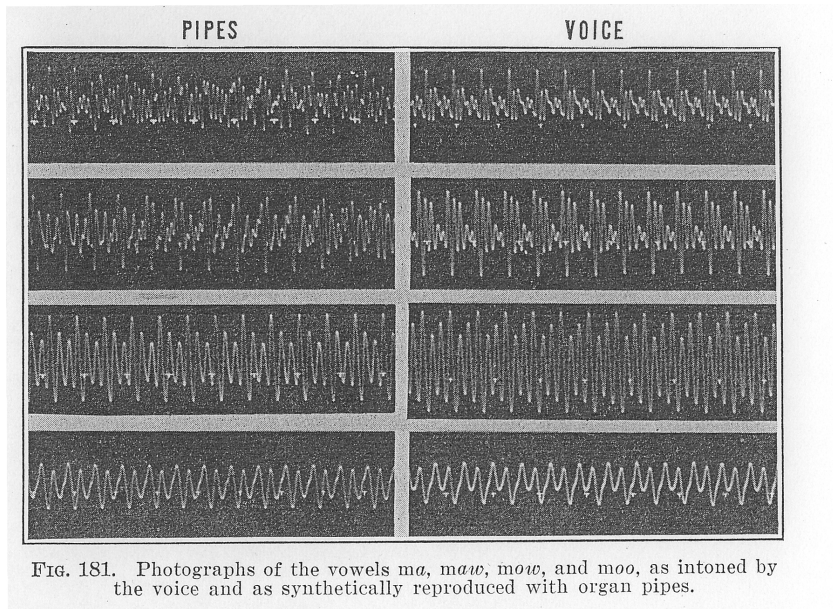
»There is no difficulty in identifying the synthetic vowels and in detecting the peculiar qualities of the voice. In this manner the general characteristics of the vowels *father* [/ɑ:/], *raw* [/ɔ:/], *no* [/o:/?], and *gloom* [/u:/], spoken by D C M [Dayton Clarence Miller], are reproduced. [...]

In the laboratory eleven vowels have been successfully reproduced by organ-pipe synthesis.« [Miller 1922, 251]

Leider werden nur drei dieser elf Vokalqualitäten explizit genannt. In den Legenden der dem Artikel beigegebenen Abbildungen werden zusätzlich die Qualitäten /æ/ und /ʊ/ genannt. Welche weiteren sechs Vokale noch synthetisierbar gewesen sein sollen, bleibt offen.

In einem nächsten Schritt versuchte MILLER, mit demselben Versuchsaufbau die spätestens seit KEMPELEN wohlbekannten »Referenzwörter« ›Mama‹ und ›Papa‹ zu synthetisieren:

»If the supply of air from the ›lungs‹ to the ›throat‹ is stopped and released by pressing the edge of the hand on the



**Abbildung 5.5:** Vergleich der Oszillogramme der vier Vokale /ɑ:/, /ɔ:/, /o:/ und /u:/ (von oben nach unten) von MILLERS Synthese mit solchen von natürlichen Sprechern [Miller 1922, 250].

rubber supply tube, the ›infant‹ is caused to cry *pa-pa* with a voice that would be considered human, if the source were unknown. [...]

If the flow of air is only partially interrupted, the sound does not altogether cease between syllables, and the pipes pronounce *ma-ma*; this is better shown by alternately producing *pa-pa* and *ma-ma*. « [Miller 1922, 253 ff.]

Wie auch das voranstehende Zitat zeigt, kommt MILLER für seine Syntheseversuche zu einem ähnlichen Ergebnis, wie es späterhin für die Repliken der KEMPELENSCHEN Sprachmaschine festgestellt wurde: Eingebettet in ein Trägerwort sind die Vokale besser erkennbar bzw. einer intendierten Qualität zuzuordnen [Miller 1922, 253].

### 5.2.2 G. Oscar Russel & Arthur Taber Jones

Der Sprachwissenschaftler G. OSCAR RUSSEL (1890–1962) berichtet von einer Beobachtung über Orgelpfeifen: Er stellte fest, dass eine metallene *labiale* Orgelpfeife mit dem Ton  $\text{dis}^2$  ( $\approx 615$  Hz) einen /a/-ähnlichen Ton gab, der sich, wenn man ihr *Labium* seitlich leicht bedeckte, in Richtung eines /u/ veränderte. Da hiermit nur eine geringfügige Tonhöhenveränderung einherging, folgerte RUSSEL, dass dieser Qualitätsunterschied durch eine andere Filterung der Teiltöne, nämlich eine Dämpfung der hohen Teiltöne, hervorgerufen worden sein musste. Der Physiker ARTHUR TABER JONES (Lebensdaten unbekannt) konnte wenige Jahre darauf diese Beobachtung bestätigen. Er konstatierte, dass sich der beobachtete Effekt mit Pfeifen höherer Tonlage (zweigestrichene Oktave und höher) besser erzielen ließ als mit tieferen [Jones 1935, 282 f.]. Leider teilen weder RUSSEL noch JONES nähere Angaben zur Bauform (*Weitenmessur*, *Labiiierung*, *Aufschnitt*) der verwendeten Pfeifen mit.

Dem Autor war es möglich, die Versuchsanordnung mit zwei in seiner Sammlung vorhandenen *Labialpfeifen* der Tonhöhe  $\text{d}^2$  ( $\approx 586$  Hz) aus Metall nachzustellen.<sup>70</sup> Beide Pfeifen besitzen eine mittlere (*prinzipalische*) *Weitenmessur*, eine *Labiiierung* von  $-\frac{1}{4,5}$  und einen *Aufschnitt* von  $-\frac{1}{4}$ . Die Pfeifen wurden mit konstantem Druck mit dem Mund angeblasen und das Signal mit *praat* aufgezeichnet.

Die »Lautqualitäten« ähneln auditiv bei unbehinderter Absprache der Pfeife in der Tat einem offenen Vokal. Die durchschnittlichen Werte für  $F_1$  von 596 bzw. 624 Hz und für  $F_2$  von 1204 bzw. 1256 Hz bewegen sich auch im Umkreis der /a/-Referenzwerte für einen männlichen Sprecher. Wurde das Labium beidseitig mit den Fingern soweit verengt, bis die Pfeife nahezu nicht mehr ansprach, ergaben sich Formantwerte von 586 bzw. 583 Hz für  $F_1$  und 1325 bzw. 1304 Hz für  $F_2$ . Diese Werte sind weit von denjenigen eines /u/ entfernt. Gleichwohl äußerte eine Person, die die Versuche des Autors zufällig mitangehört hatte, sie habe ein deutliches /u/ gehört.

Die mit der Labiumsverengung zusammenhängende Tonhöhenerniedrigung war indessen durchaus merklich ( $\approx -30$  Hz), was zwar RUSSELS Feststellung widerspricht, aber aus orgelbaulicher Sicht als vollkommen natürlich bezeichnet werden muss.

<sup>70</sup>Vgl. die Tonbeispiele auf der CD-R im Appendix (S. LXXII).



## 5.3 Vokalsynthese mit Vokaltraktmodellen

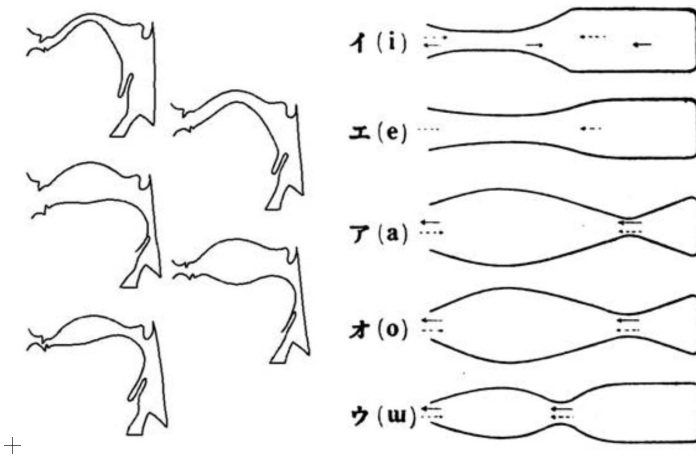
In den 1930er Jahren modellierten die japanischen Forscher TSUTOMU CHIBA und MASATO KAJIYAMA nach dem Vorbild von Sagittalschnitten dreidimensionale Tonmodelle des menschlichen Vokaltraktes für die fünf japanischen Vokale /i, e, a, o, u/ (Abb. 5.6), mit denen bereits überzeugend klingende synthetische Laute erzeugt werden konnten [Chiba 1941]. Diese Versuche wurden 60 Jahre später von TAKAYUKI ARAI hauptsächlich zu Unterrichtszwecken wiederholt und umfassend zu einer ganzen Serie von anatomisch basierten Vokalsynthesen weiterentwickelt. Hierzu übertrug er die tönernen Vokaltraktmodelle von CHIBA & KAJIYAMA zunächst auf Plexiglas (Abb. 5.7 links) und konnte hiermit die früheren Ergebnisse bestätigen. Beide Modelle wie auch ein drittes, aus diesen abgeleitetes und vereinfachtes, stellen jedoch den menschlichen Vokaltrakt als ein gerades Rohr mit wechselndem Querschnitt dar. Dies ist in sofern ungenau, als es sich in der anatomischen Realität um ein annähernd rechtwinklig gekrümmtes Rohr handelt (Abb. 112).

ARAI entwickelte daher ein Vokaltraktmodell, das sich stark an einem anatomisch korrekten Sagittalschnitt orientiert und aus jeweils fünf verschieden starken Plexiglasplatten zusammengesetzt ist. Hiermit lassen sich die tatsächlichen artikulatorischen Zusammenhänge weitaus anschaulicher machen als mit den bereits recht abstrakten zylindrischen Modellen. Auch dieses war jedoch statisch; Artikulationsprozesse ließen sich nicht nachvollziehen. Hieraus entstanden schließlich Modelle mit beweglichem Zungen- bzw. Velumsanalogon [Arai Lab 2013].

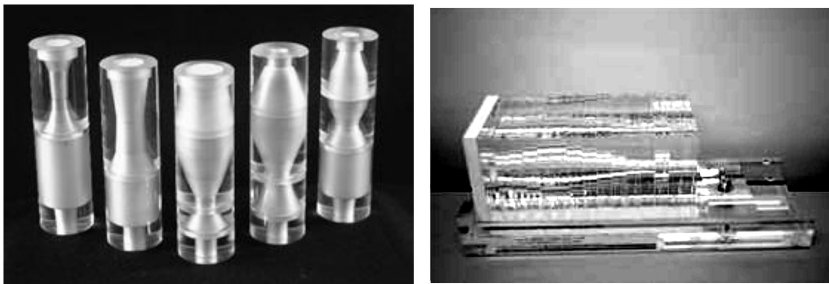
Ein abweichendes Konzept der Modellierung von Vokaltrakten entwickelten die ebenfalls japanischen Forscher NORIKO UMEDA und TERANISHI. Es handelt sich ähnlich einem frühen Modell ARAIS (Abb. 5.7 rechts) um einen zunächst rein zylindrischen geraden Vokaltrakt, dessen Querschnitt durch seitlich eingeführte »Schieber« unterschiedlicher Dicke modifiziert werden kann (Abb. 5.8). Durch eine computergestützte Steuerung dieser Schieber ist nicht nur die Modellierung von Monophthongen, sondern auch die von Diph- oder gar Triphthongen möglich. Die Anregung geschieht durch das Signal eines Lautsprechers.

## 5.4 »Sprechende Köpfe« im 21. Jahrhundert

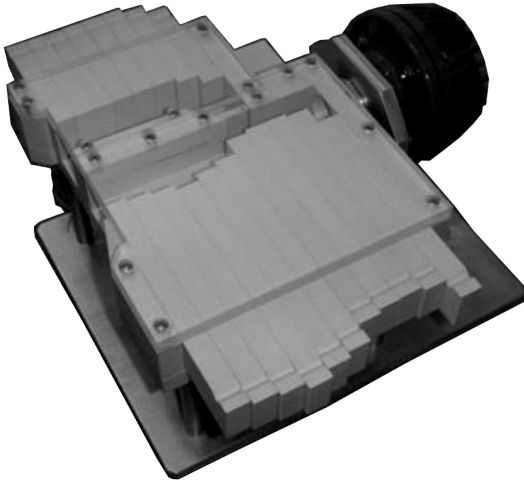
Auch 150 Jahre nach den BELL-Brüdern und sowohl wissenschaftlich als auch philosophisch geradezu Lichtjahre von den Versuchen eines Robert BACON oder ALBERTUS MAGNUS entfernt wird das Konzept des »Spre-



**Abbildung 5.6:** Schematische Schnitzzeichnungen des menschlichen Vokaltrakts und der daraus von CHIBA abgeleiteten Modelle [Chiba 1941, 144].



**Abbildung 5.7:** Die statischen Vokaltraktmodelle von ARAI nach CHIBA & KAJIYAMA. Links: Vokaltrakte für /i, e, a, o, u/ (von l. n. r., hier oben = Mundöffnung), rechts: Scheibenmodell für /a/ [Arai Lab 2013].



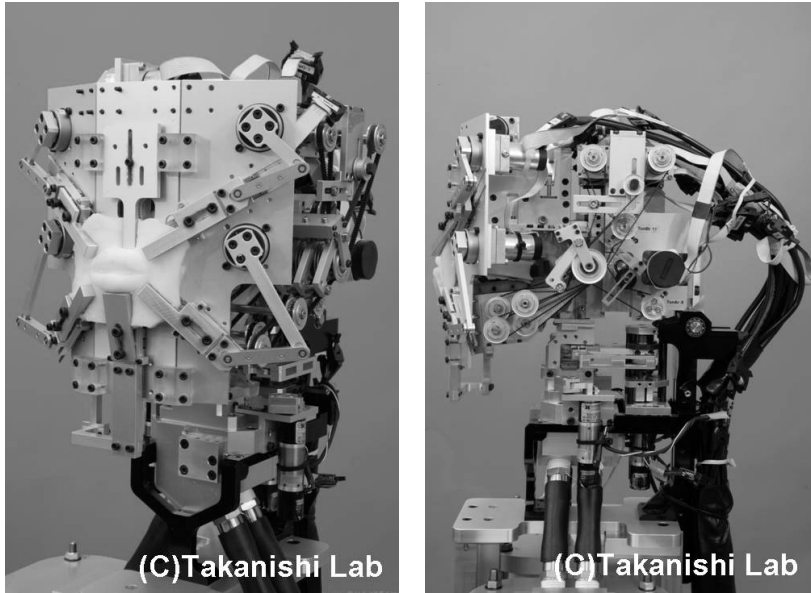
**Abbildung 5.8:** *Das Vokaltraktmodell von UMEDA & TERANISHI [nach Arai Lab 2013].*

chenden Kopfes« von mehreren Seiten erneut aufgegriffen. Während die »Waseda Talkers« hierbei das Prinzip der mechanischen Sprachsynthese auf ein technologisch vollkommen neues Niveau heben, stellt der »Furhan Talking Head« keine Sprachsynthese im engeren Sinne mehr da, erst recht keine mechanische. Trotzdem spiegelt sich in ihm die alte Idee eines autark und spontan kommunizierenden humanoiden Apparates deutlich wider.

#### 5.4.1 Die »Waseda Talkers«

Eine japanische Forschergruppe um KATARO FUKUI verfolgt seit einigen Jahren mit dem »Waseda Talker« einen hinsichtlich der Anregung den thematischen Rahmen dieser Arbeit sprengenden Ansatz zur mechanischen Sprachsynthese. Da es sich gleichwohl um ein mechanisches Synthesekonzept handelt, sei es hier knapp beschrieben. Die Ausführungen beziehen sich hierbei schwerpunktmäßig auf die Angaben der offiziellen Internetpräsenz des Projektes [Anonymus 2013b].

Ziel dieser Forschungen war es zunächst, ein sowohl anatomisch als auch materialtechnisch dem menschlichen Vorbild so weit als irgend möglich entsprechendes Analogon zu schaffen, an dem sich die Phonations- und Artikulationsvorgänge noch besser als beim Menschen in vivo erforschen



**Abbildung 5.9:** Der »Waseda Talker No. 7« [Anonymus 2013b].

lassen [Fukui et al. 2008, 341]. Auf der Projektwebsite wird die Motivation entsprechend wie folgt zusammengefasst:

»This basic theory [der menschlichen Stimmgebung] has been clarified; however, the control of the complex movement of speech organs and detailed phonemes in speech production are still unclear.

For clarifying human speech, mechanical models are superior method. [...] From few centuries ago, peoples developed several mechanical models of speech organs, however, these are simple mechanism and difficult to simulate the movement.« [Anonymus 2013b, Introduction]

Hieraus entstand seit 2000 eine mittlerweile acht Prototypen umfassende Serie von humanoiden, sprechenden Robotern, den »Waseda Talkers«, die eine wohl einmalige Kombination von humanoider robotischer Technologie und Sprachsynthese darstellen. Ab dem Modell Nr. 5 wurden diese Konstruktionen auch mit einem dreidimensionalen Modell der menschlichen Glottis und des Vokaltraktes ausgestattet. Als Material für den gesamten Mechanismus wird der flexible Kunststoff »Septon« verwendet,

der der biomechanischen Struktur des menschlichen Gewebes nahekommmt [Fukui et al. 2008, 341].

Wurde beim »Waseda Talker No. 5« noch ein sehr eng an den Menschen angelehnter dreidimensional durchgestalteter Mechanismus für die Stimmlippen benutzt, verwendet der »Waseda Talker No. 7« hier schon eine Weiterentwicklung, die sich allerdings anatomisch vom Vorbild entfernt: Hier wird ein rotierendes Scheibenpaar eingesetzt, das direkt auf den Stimmlippen sitzt und deren Spannung exakt reguliert [Fukui et al. 2008, 341].

Das jüngste Modell Nr. 7 verfügt über eine mechanische Analogie der Lungen, der Stimmlippen, der Zunge, des Unter- und Oberkiefers, des Velums, des Nasenraums und der Lippen nach dem Vorbild eines männlichen Modells. Die Lungen werden durch zwei Kolbenzylinder repräsentiert, deren Austrittsmenge und Druck variiert werden kann.

Als besonders interessant müssen die Lösungen für Zunge und Lippen bezeichnet werden, waren dies doch in allen früheren Versuchen stets die Probleme, an denen die Forscher scheiterten. Beim »Waseda Talker No. 7« ist die Zunge dreidimensional einem menschlichen Vorbild nachmodelliert und in den drei für die Sprachproduktion zentralen Bereichen der Zunge (Spitze, Blatt und Rücken) separat und authentisch steuerbar. Die Lippen sind aus einem weichen Kunststoff gebildet und verfügen über insgesamt fünf mechanische »Muskeln«, über die sie in die notwendigen vorbildgerechten Positionen für die momentan möglichen Vokale /i, e, a, o, u/ verschoben werden können. Zunge und Lippen können über zwei verschiedene Methoden für die jeweiligen Laute angesteuert werden. Zum einen besteht die Möglichkeit, auf bereits vorprogrammierte Parameter für jedes Einzelphonem zurückzugreifen. Diese enthalten jedoch jeweils nur Werte für die stabilen Phasen des jeweiligen Lautes und werden in zusammenhängenden Äußerungen über Sinusfunktionen miteinander verknüpft. Zum anderen besteht die Möglichkeit, auf artikulatorische Daten zurückzugreifen, die via elektromagnetischer Artikulografie (EMA) erhoben wurden. Die Artikulationsvorgänge und damit auch das akustische Ergebnis unterscheiden sich erheblich voneinander.<sup>71</sup>

Darüber hinaus ist es möglich, verschiedene Stimmqualitäten darzustellen: Neben der normalen Sprechstimme verfügt diese Synthese auch über Repräsentationen von behauchter und gepresster Stimme [Fukui et al. 2008, 342].

---

<sup>71</sup>Vgl. die Beispielsynthesen unter <http://www.takanishi.mech.waseda.ac.jp/top/research/voice/index.htm#3rd> (zuletzt gesichtet am 27.06.2013)

### 5.4.2 Der »Furhat Social Companion Talking Head«

Auch das zweite moderne Konzept eines »Sprechenden Kopfes« würde bei detaillierter Darstellung den thematischen Rahmen dieser Arbeit sprengen. Die knappe Darstellung folgt im Wesentlichen AL MOUBAYED et al. (2013).

Bei diesem Projekt des Departement of Speech, Music and Hearing des KTH Royal Institute of Technology (Stockholm) handelt es sich letzten Endes um einen dreidimensionalen Avatar, der eine ganze Reihe von hochentwickelten Simulationslösungen zu einem durchaus beeindruckenden Gesamtkonzept verbindet. Neben einer projizierten Mimik verfügt »Furhat« auch über eine Gesichtserkennung und eine anspruchsvolle Dialogsoftware, also Spracherkennung und Spracherzeugung. Das intendierte Konzept eines humanoiden »Gefährten« und kommunikativen Konterparts impliziert hierbei nicht nur eine »intelligente« Sprachsynthese von hoher Qualität, sondern auch eine hochentwickelte Spracherkennung. Angesichts dieses anspruchsvollen Designs wirkt die ebenfalls vorhandene Gesichtserkennung beinahe als zweitrangige Dreingabe.

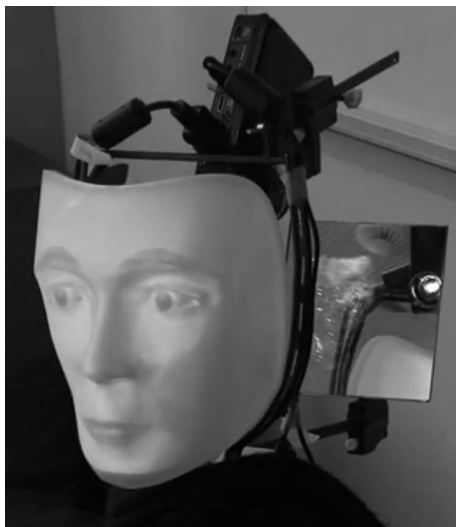
Die »Mimik« des »Furhat« geschieht über einen Mikroprojektor, der bewegliche Repräsentationen für Augen und Mund von innen auf eine dreidimensionale Gesichtsmaske projiziert. Die Animation insbesondere des Mundes wird hierbei mit der Sprachausgabe synchronisiert. Anders als bei zweidimensional auf Bildschirmen dargestellten Avataren (wie bspw. von WAN et al. (2013) beschrieben) oder mechatronischen Mimiksimulationen ist durch die Projizierung auf ein dreidimensionales Kopfmodell eine sehr dynamische und »fließende« Simulation von Mimik möglich. Eine Motorik im Nackenbereich erlaubt es dem »Furhat« zudem, sich seinem Gesprächspartner zuzuwenden und seine Äußerungen ggf. durch Kopfbewegungen zu unterstreichen. Dies ist besonders wesentlich, da das System mit mehreren Personen zugleich interagieren kann.

Die Sprachausgabe nutzt die Synthesesoftware »CereVoice« des Entwicklers CereProc, die eine verhältnismäßig natürlich klingende und prosodisch relativ vielfältige Synthese ermöglicht.

## 5.5 Exkurs: Ein »sprechendes Klavier«

Eine gänzlich andere Form der Sprachsynthese (im erweiterten Sinne) erprobten der Komponist PETER ABLINGER (\*1959) und WINFRIED RITSCH (\*1964) 2008 im Rahmen des Festivals »Wien modern«. Genau genommen handelt es sich hierbei um eine Zwischenform zwischen Musik und Sprache.

Wie in anderen hier vorgestellten modernen mechanischen Sprachsyntheseansätzen geschieht auch hierbei die Steuerung letztlich über einen



**Abbildung 5.10:** *Der »Furhat« ohne die namensgebende Kopfbedeckung. Zu erkennen ist die sonst verborgene Projektionstechnik für die Mimik. [Anonymus 2013c].*

Computer. Die Synthese an sich wird jedoch von einem Klavier erzeugt, das bis auf einen elektromagnetischen Aufsatz zur Steuerung der Tasten sich in nichts von einem »normalen« Klavier unterscheidet. Als Ausgangsmaterial dienen die von einem Kind gelesenen Texte, wie beispielsweise die »Declaration of the International Environmental Criminal Court«, die aufgenommen werden. Anschließend wird das Frequenzspektrum dieser Aufnahmen in eine »Partitur« umgerechnet, nach der die Hämmer des Klavieres elektronisch angesteuert werden. Durch das hochpräzise koordinierte gleichzeitige Anschlagen jeweils einer ganzen Reihe von Tönen unterschiedlicher Tonhöhe wird ein entfernt an menschliche Sprache erinnerndes Frequenzband erzeugt. Mit dieser Methode werden selbst stimmlose Plosive simuliert. Eine Reihe von Demonstrationen bzw. Kompositionen findet sich bei [ABLINGER 2013].

Die effektive Verständlichkeit dieser Synthese stellt ein interessantes psycholinguistisches Phänomen dar: Faktisch dürfte sie noch hinter den Konzeptionen MICALS und KEMPELENS zurückbleiben, jedoch sorgen in den Videoinstallationen stets eingeblendete Untertitel dafür, dass die Zuhörer ein Gefühl des souveränen »Verstehens« entwickeln. Einem Blindversuch hält die Synthese indessen nicht stand: Mehrere Kollegen des Autors,

die ohne jede Vorinformation mit der synthetisierten Version der »Declaration« konfrontiert wurden, konnten das Gehörte nicht spontan als Sprache oder sprachenähnlich identifizieren.

## 5.6 Zusammenfassung

Das 20. Jahrhundert war bestimmt durch vollständig neue Ansätze auf dem Gebiet der Forschung zur mechanischen Sprachsynthese. Einen nicht zu unterschätzenden Einfluss hatte hier – wie auch auf zahlreiche andere Gebiete – die nunmehr beherrschbare Elektrizität, die eine Vielzahl neuer Technologien hervorbrachte. Zu den bekanntesten Exponenten der damaligen Sprachsyntheseforschung gehört zweifellos der »Voder« HOMER DUDLEYS. Dass er sich in der Folge nicht durchsetzen konnte, dürfte vor allem daran gelegen haben, dass es auch für ihn noch – wie schon für die Sprachmaschine KEMPELENS 150 Jahre zuvor – unerlässlich war, seine Bedienung wie ein Musikinstrument zu erlernen und für eine überzeugende Synthese virtuos zu beherrschen.

Aber auch auf dem Gebiet der »klassischen«, also mechanischen Sprachsynthese, die durch die Nachbildung der menschlichen Anatomie zu einer befriedigenden Synthese zu gelangen hoffte, wurden zahlreiche neue Konzepte entwickelt, die nur noch sehr begrenzt auf die Entwicklungen des 18. und 19. Jahrhunderts Bezug nahmen. Zu nennen sind hier die Vokaltraktmodelle von RIESZ, die – quasi als Reminiszenz – als Schallquelle auf *Zungenpfeifen* zurückgriffen. Die erheblich späteren »Waseda Talkers« stellen den bislang jüngsten Versuch dar, mithilfe vorbildgetreuer Modellierung des Vokaltraktes wie auch der Artikulationsorgane zu einer überzeugenden Synthese zu gelangen.

Die Vokaltraktmodelle der japanischen Forscher CHIBA & KAJIYAMA sowie UMEDA & TERANISHI wurden zwar insbesondere zu Unterrichts- und Anschauungszwecken entwickelt, stellen aber gleichwohl ein junges Beispiel des Bemühens um eine exakte Nachvollziehung der menschlichen Vokalarthikulation und somit eine späte Fortentwicklung der Arbeiten KRATZENSTEINS dar.

Mit dem »Furhat Social Companion Talking Head« ist schließlich eine Brücke geschlagen zwischen dem »sprechenden Kopf« als wohl ältestem äußeren Erscheinungsbild einer Sprachsynthese überhaupt und einem vollkommen neuartigen und erst jetzt durch den Einsatz von sehr leistungsfähigen Computern möglich gewordenen Synthesekonzept, das erstmals der spontanen Interaktion zwischen Sprachsynthese und menschlichem »Gesprächspartner« den höchsten Stellenwert einräumt.



Die vorangegangenen Kapitel konnten aufzeigen, welche Entwicklung die Sprachsyntheseforschung seit ihrem scheinbaren Beginn bereits in der Antike genommen hat. Von seriösen (lies: nicht auf Täuschung beruhenden) Sprachsynthesekonzepten kann indessen erst frühestens ab dem 17. Jahrhundert gesprochen werden. Abbildung 5.11 zeigt synoptisch alle in dieser Arbeit behandelten Sprachsynthesekonzepte. Die zeitliche Ballung am Ende des 18. Jahrhunderts und in der Mitte des 19. Jahrhunderts wird hierbei überdeutlich.

Über eine Zeitspanne von gut 800 Jahren – seit den legendenhaften Konstruktionen ROBERT BACONS (1214–1292) und ALBERTUS MAGNUS' (-1200–1280) – wurde die Assoziation von Sprachsynthese in den meisten Fällen mit der äußerlichen Erscheinung eines »sprechenden Kopfes« verknüpft. Dabei sind die Gründe hierfür zumindest grundsätzlich dieselben geblieben: Sowohl ATHANASIUS KIRCHER (1602–1680) als auch noch FUKUI et al. äußerten die Überzeugung, dass eine naturgetreue Nachbildung der menschlichen Anatomie Voraussetzung für eine möglichst authentische Synthese sein müsse. Einziger Unterschied hierbei ist, dass KIRCHER und seine Zeitgenossen (und Vorläufer) ihre Versuche unter sehr eingeschränkten Voraussetzungen durchführen mussten, sowohl, was das anatomisch-artikulatorische Wissen betrifft als auch, was die technologischen Möglichkeiten angeht. Vor dem Hintergrund unserer heutigen Selbstwahrnehmung als hochtechnologisierte Gesellschaft erscheint es bemerkenswert, dass auch die »Waseda Talkers« bislang ein qualitativ nur sehr mäßiges und hinsichtlich ihres Repertoires erheblich eingeschränktes Syntheseresultat zu liefern im Stande sind.

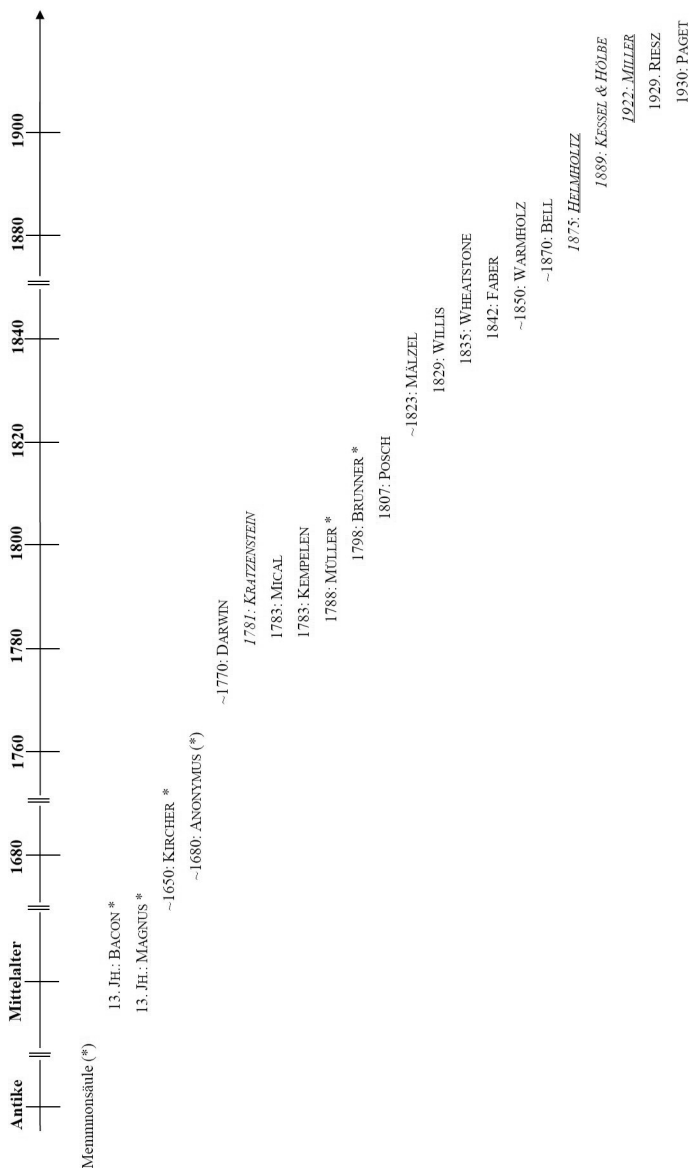
Die (starke) Beschränkung des Syntheserepertoires scheint – ähnlich wie die *Idée fixe* des »sprechenden Kopfes« – ein ebenfalls beständiges Thema in der Forschung zur mechanischen Sprachsynthese zu sein. Dies ist aus phonetischer Sicht wenig verwunderlich, erfordern die menschlichen Sprachlaute doch Artikulationsvorgänge höchst unterschiedlicher Komplexität. Dass für eine ganze Reihe von historischen Sprachsyntheseprototypen die »frühkindlichen« Äußerungen »Mama« und »Papa« überliefert sind, überrascht beim Wissen um die hierzu notwendigen Artikulationsbewegungen wenig. WOLFGANG VON KEMPELEN (734–1804) überliefert für seine Sprachmaschine recht komplexe Äußerungen wie »Leopoldus Secundus – Romanorum Imperator – Semper Augustus« [Kempelen 1791a, 456]. Wie diese zustande gekommen sein könnten, muss offen bleiben, da KEMPELENS Konstruktion doch ebenfalls an der zentralen Krux aller mechanischen Sprachsynthesekonzepte krankte, dem Fehlen eines (adäquaten) Zungenanalogons.

Es konnte aufgezeigt werden, wie beherrschend zumindest im 18. und 19. Jahrhundert die Vorstellung war, in den *Vox humana-Registern* (VH) der Orgeln den rudimentären Prototyp einer mechanischen Sprachsynthese zu besitzen. Eine besonders exponierte Stellung in diesem Themenkomplex kommt CHRISTIAN GOTTLIEB KRATZENSTEIN (1723-1795) und seinem Mentor LEONHARD EULER (1707-1783) zu. Ihre Arbeiten bezogen sich noch expliziter als andere auf das Orgelregister VH bzw. die Orgel an sich als postulierte Vorform der Sprachsynthese. Zudem wirkten die Ergebnisse ihrer Bemühungen auch zurück auf das Gebiet des Orgelbaues, indem dort die (anscheinend) von KRATZENSTEIN oder einem Orgelbauer in seiner engeren Umgebung entwickelten *durchschlagenden Zungen* vom Orgelbau des 19. Jahrhunderts, besonders in Deutschland, sehr rasch und vielfältig aufgegriffen wurden und seine Klangästhetik entscheidend prägten.

Aus heutiger Sicht scheint die Vorstellung, es könne sich bei der VH zumindest um die Vorform einer Vokalsynthese handeln, beinahe absurd, doch macht es nachdenklich, wenn man sich vor Augen führt, mit welcher abgeklärter und kritischer Sicht dieselben Forscher, die sich auf die VH versteift hatten, andere Aspekte der Sprachforschung und andere Themenbereiche abhandelten. Warum also übte ein Orgelregister, dessen Klangfarbe wir Heutigen wohl zumeist mit »schnarrend«, »quäkend« oder jedenfalls wenig menschenähnlich beschreiben würden, eine so große Faszination auf gestandene und aufgeklärte Forscher jener Zeit aus? War es eine Frage der Wahrnehmung, klangen die *Register* damals vollkommen anders, als wir sie heute kennen oder nahm man ihren Klang anders wahr? Spielte eine mittlerweile verlorene obertonlastige Gesangstechnik ähnlich dem *Canto a tenore* hierbei eine Rolle?

Im folgenden zweiten Teil dieser Arbeit soll zunächst auf experimenteller Grundlage der Frage nachgegangen werden, ob sich grundsätzliche Ähnlichkeiten oder gar Übereinstimmungen zwischen dem Klang einer VH und menschlicher Sprache nachweisen lassen. Sind an heute noch erhaltenen VH-*Registern* des 18. Jahrhunderts akustische Faktoren nachweisbar, die das Postulat einer Ähnlichkeit zur menschlichen Stimme stützen können?

In den anschließenden Kapiteln werden Fragen zur Rekonstruktion bzw. möglichen Weiterentwicklung der Sprachsynthesekonzepte CHRISTIAN GOTTLIEB KRATZENSTEINS und WOLFGANG VON KEMPELENS behandelt. Insbesondere für die Sprachmaschine KEMPELENS existieren sowohl aus seiner Hand wie auch aus späteren Quellen zahlreiche Vorschläge, die hauptsächlich auf die Optimierung der Schallquelle abzielen.



**Abbildung 5.11:** Synoptischer Zeitstrahl der mechanischen Sprachsynthesekonzepte von der Antike bis zum Beginn des 20. Jahrhunderts (\*: Pseudo-Synthesen, (\*): Vermutliche Pseudosynthesen, kursiv: Vokalsynthesen, kursiv: Vokalsynthesen mit Labialpfeifen).



## Teil II

# Historische Sprachsynthesen im Experiment



# Kapitel 6

## Vox humana – Eine Vokalsynthese?

Die im Teil I geschilderten historischen Wechselwirkungen zwischen dem Orgelregister *Vox humana* (VH)<sup>72</sup> – bzw. dem Instrument Orgel im Allgemeinen – und den frühen Versuchen zur Sprachsynthese wirken auf uns Heutige einigermaßen befremdlich. Obwohl eine gewisse Ähnlichkeit zwischen der mutmaßlich urwüchsigen Gesangspraxis des *Canto a tenore* und dem Klang historischer VH besteht, kann aus dieser Koinzidenz kein Beweis dafür abgeleitet werden, dass die VH im obertönigen Kehlgesang ihr klangliches Vorbild und den Ursprung ihres Namens hat und dadurch zur *Idée fixe* der Sprachforscher des 17. bis 19. Jahrhunderts wurde.

Eine phonetische Analyse von VH-Tönen sollte hierzu mehr Klarheit bringen. Dafür wurde der Klang mehrerer VH sowohl mit instrumentalphonetischen Methoden als auch durch eine Reihe von Perzeptionsexperimenten auf eine Sprach- bzw. Vokalähnlichkeit hin überprüft.

### 6.1 Frühere akustische Analysen

WERNER LOTTERMOSER (1909–1997, deutscher Akustiker) geht in seinem Kompendium über die Akustik von Orgeln knapp auf den speziellen Klangaufbau der VH ein:

---

<sup>72</sup>Für die im Folgenden kursiv gesetzten *orgelbaulichen Fachbegriffe* vgl. das Glossar ab S. LXII.

»Viele Register [Lingualregister], vor allem die Zungenpfeifen der Vox humana [...] besitzen in ihren Klangspektren auffallende Formanten, die durch die obertonreichen Klänge der Zungensteuerung und die Hohlräume der Schallbecher zustande kommen. So ist es sinnvoll, deren Klänge direkt mit denen der Vokale zu vergleichen und ihre Klangfarbe durch diejenige der entsprechenden Laute zu beschreiben. [...] Wenn man ein Vox-humana-Register zieht [einschaltet], so erzeugt man damit wohl einen vokalähnlichen Klang mit Formanten, doch sind die FS [Frequenzschwankungen] viel kleiner als bei der menschlichen Stimme. Die Ähnlichkeit wird aber größer, wenn man die Vox-humana [sic!] mit dem Tremulanten betreibt, durch welchen eine regelmäßige Amplituden- und Frequenzmodulation erzeugt wird. Durch diese wird die Vokalähnlichkeit der Vox-humana, wie man sich leicht überzeugen kann, auffallend größer.«<sup>73</sup> [Lottermoser 1983a, 59 und 62]

Knapper und zugleich präziser beschrieb er bereits knapp 50 Jahre zuvor dasselbe Phänomen:

»Der entstehende Hohlraum [der teilgedeckte Resonator der VH, der bei allen Tönen ähnliche Abmessungen aufweist] weist für Pfeifen tiefer und hoher Tonlage einen Eigenton konstanter Frequenz auf, und es entstehen Formantgebiete ähnlich denen der menschlichen Vokale.« [Lottermoser 1936, 15]

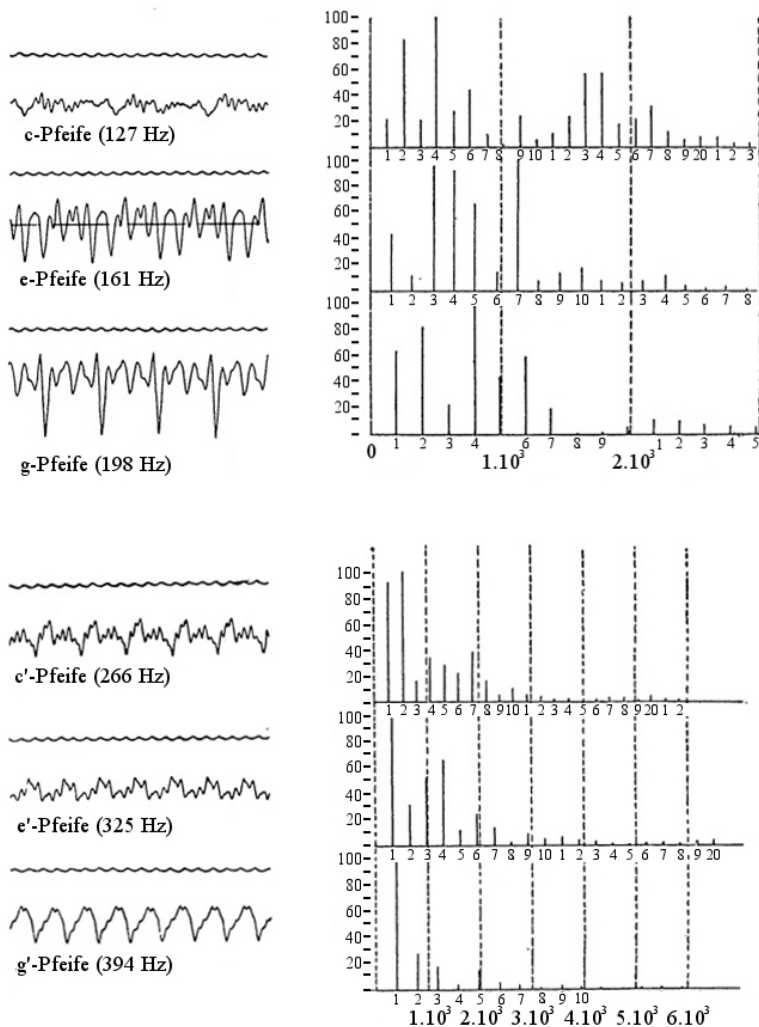
Er bringt auch einen spektralen Vergleich zwischen einer VH aus dem Jahre 1626 und einer modernen Konstruktion vom Ende der 1920er Jahre (Abb. 6.1). Den modernen VH-Pfeifen attestiert er hierbei einen Ä-ähnlichen Klang mit Formantgebieten um 978 Hz und 1958 Hz.<sup>74</sup> Bei den Pfeifen von 1626 seien hingegen keinerlei Formantstrukturen erkennbar, was mit deren kontinuierlich anwachsender Resonatorgröße zusammenhänge (die qua definition eigentlich untypisch für eine VH ist). [Lottermoser 1936, 48 ff.]

ADELUNG erwähnt, dass der Naturton von *Lingualregistern* allgemein sehr obertonreich sei, da man bis zu 60 Teiltöne feststellen könne. Durch

<sup>73</sup>Von zentraler Bedeutung ist hierbei allerdings, welche Art von Tremolo der *Tremulant* erzeugt. Sind die Bebungen von ihrer Frequenz oder Intensität denen der menschlichen Singstimme nicht ähnlich, kann sich auch die Suggestion einer solchen nicht einstellen.

<sup>74</sup>Letzterer Wert entspricht in der Tat dem weiblichen Durchschnittswert für F<sub>2</sub> bei /ε/ (F<sub>1</sub> 595 Hz) [Simpson 1998].





**Abbildung 6.1:** Oszillogramme (links) und Spektren (rechts) von Pfeifen einer historischen (oben) und einer modernen Vox humana (unten) [nach Lottermoser 1936, 48 f.]. Die Skalierung der Spektren bezeichnet auf der x-Achse die Obertöne, auf der y-Achse deren lineare Intensität.

die geringe Filterung der kurzen *Resonatoren* blieben bei *Regalen* hiervon die meisten im Klang erhalten, während die langen *Resonatoren* beispielsweise von *Trompetenregistern* die meisten hohen Teiltöne eliminieren. [Adelung 1982, 108]

Das Charakteristikum aller Bauformen der VH ist also, dass die Proportionen des Resonators über weite Bereiche des Klaviaturambitus annähernd konstant bleiben und somit eine ähnliche Konfiguration erzeugen wie es bei unterschiedlich hoch gesungenen Vokalen beim Menschen der Fall ist.

## 6.2 Eigene akustische Analysen

### 6.2.1 Datengrundlage

Die Reihe von Zitaten aus knapp 300 Jahren in Kap. 2 und im vorangegangenen Abschnitt zeigt die scheinbare Vokalähnlichkeit der VH und stellt sie in Beziehung zu einer prototypischen Sprachsynthese. Objektive Belege, also Messungen zu dieser Behauptung fehlen allerdings bislang bis auf die referierten Untersuchungen LOTTERMOSERS völlig. Diese können jedoch schon allein aufgrund ihrer Entstehungszeit modernen Anforderungen an präzise akustische Messungen nicht genügen.

Gerade im hier hergestellten Kontext – und nicht zuletzt vor dem Hintergrund, dass sich bereits bei früheren Forschungen des Autors historische »Tatsachenberichte« als oftmals recht blumig und unpräzise erwiesen hatten – erschien es zudem als wünschenswert, wenn nicht gar notwendig, diese Quellen, wie z. B. bei GRESS (2007, 27) angeführt, nicht blind zu zitieren, sondern einer durchaus kritischen Überprüfung zu unterziehen. Dies bedeutet konkret, dass die angeführten Literaturquellen durch praktische Untersuchungen und Messungen zu überprüfen und entweder zu verifizieren oder falsifizieren waren: Weisen *VH-Register* des 18. Jahrhunderts charakteristische Eigenschaften (gesungener) Vokale auf?

Um dies messtechnisch zu untersuchen, war es notwendig, einige historische *VH-Register*, die in ihren wesentlichen (klangbestimmenden) Teilen noch aus dem 18. Jahrhundert (und damit aus der gleichen Epoche wie die meisten o. g. Schriftquellen) stammen, nach genau definierten Kriterien akustisch zu vermessen. Anhand eines Vergleichs der jeweiligen Spektren bzw. Spektrogramme mit denjenigen von menschlichen Vokalen soll dann aufgezeigt werden, in wie weit alle, einige oder eben gar keine dieser vermessenen *Register* den Formantstrukturen statischer Vokale nahekommen.

Bei der Konzeption einer solchen Versuchsreihe stellt sich indessen ein ganz grundsätzliches Problem: Orgelpfeifen allgemein, insbesondere aber *Zungenregister*, unterliegen mittel- und erst recht langfristig dem Einfluss starker klangästhetischer Veränderungen. So stieß der Orgelklang, wie er im 18. Jahrhundert geschätzt wurde, im darauffolgenden Zeitalter der Romantik und Symphonik bei Organisten, Komponisten und auch Orgelbauern auf Unverständnis. Wurden Orgeln (zu dieser und anderen Zeiten) nicht vollständig neu gebaut, unterzog man sie großteils zumindest einem tiefgreifenden klanglichen Umbau. Diesem fielen in der Romantik besonders häufig die *Zungenregister* vorangegangener Epochen gänzlich zum Opfer und dies besonders dann, wenn es sich – wie bei der VH – um *kurzbecherige Linguale* handelte. Wo diese nicht kurzerhand beseitigt wurden, suchte man sie durch Austausch der *Zungenblätter* und/oder andere tiefgreifende intonatorische Maßnahmen dem aktuellen Geschmack anzunähern. Noch bis in die siebziger Jahre des 20. Jahrhunderts wurden historische Orgelregister dem aktuellen Zeitgeschmack oft irreversibel angepasst, trotz eines bereits seit mehreren Jahrzehnten existierenden Bewusstseins für die grundsätzliche Denkmalswürdigkeit historischer Instrumente. Hinzu kommt, dass insbesondere die *Zungenblätter* auch unabhängig von solchen willkürlichen Eingriffen recht hohen mechanischen Belastungen ausgesetzt sind, was langfristig ihre kristalline Struktur und damit auch den von ihnen erzeugten Klang merklich und ebenfalls irreversibel verändert.

Daher muss festgestellt werden, dass kein einziges historisches *Zungenregister* in seiner klanglichen Originalgestalt überliefert ist (d. h. vollständige Erhaltung aller originalen Bauteile in ihrer ursprünglichen klanglichen Konstitution). Dies gilt allerdings – mit nur kleinen Einschränkungen – auch für sämtliche anderen, *labialen Orgelregister*.<sup>75</sup> Allerdings hat die moderne organologische Forschung inzwischen umfassende Erkenntnisse über die Fertigung und Intonation historischer Orgelpfeifen gewonnen, so dass man bei einigen weitgehend erhaltenen und gewissenhaft restaurierten *Zungenregistern* davon ausgehen kann, dass sie in ihrer restaurierten Form (z. B. mit originalgetreu ersetzten oder restaurierten *Zungenblättern*) der ursprünglichen klanglichen Intention zumindest nahe kommen.

Bei der Auffindung von VH-*Registern*, auf die diese Bedingungen zutrafen, konnte die umfangreiche organologische Bibliothek des Autors wertvolle Hilfe leisten. Erhaltene historische Orgeln des 18. Jahrhunderts im näheren Umkreis des Wohnortes des Autors wurden dahingehend überprüft, ob sie eine VH besitzen und in wie weit diese noch in ihren Ori-

<sup>75</sup>Hierfür verantwortlich ist u. a. die langfristige Veränderung der kristallinen Struktur von Metall (z. B. durch Patina). Zudem werden auch labiale Pfeifen im Laufe von Jahrzehnten in aller Regel intonatorisch verändert.

nalteilen existiert. Drei Instrumente der renommierten Orgelbauerfamilie STUMM wurden ausgewählt, deren jeweilige VH nach aktueller Quellenlage noch als authentisch erhalten (d.h. mit erhaltenem oder originalgetreu ersetztem *Zungenblatt*) angesehen werden kann:<sup>76</sup>

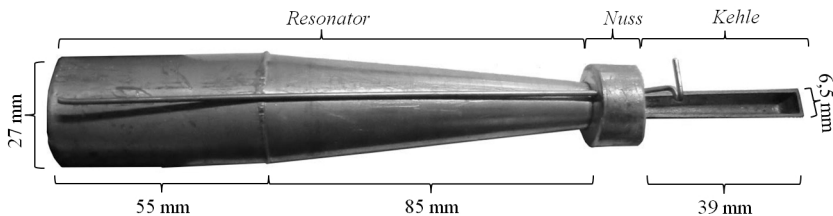
- Ehem. Abteikirche Amorbach/Odenwald (JOH. PHILIPP & JOH. HEINRICH STUMM, 1782) [Theobald 1983, 190]<sup>77</sup>
- Ev. Schlosskirche Meisenheim/Glan (JOH. PHILIPP & JOH. HEINRICH STUMM, 1767) [Adolph 2009, 14]
- Ev. Stephanskirche Simmern/Hunsrück (JOH. PHILIPP & JOH. HEINRICH STUMM, 1782) [Anonymus 2009, 123]

Um LOTTERMOSERS Analysen mit modernen Methoden nachzuvollziehen, wäre es sehr reizvoll gewesen, ergänzend zu den o. g. drei *Registern* auch Aufnahmen der sagenumwobenen *Vox humana* der weltberühmten Orgel der ehemaligen Benediktinerabtei Weingarten (JOSEPH GABLER 1737/50) zu machen (vgl. Zitat auf S. 18). Leider sah man sich auf entsprechende Nachfrage seitens der dortigen Kirchengemeinde nicht in der Lage, die Orgel für etwas anderes als den liturgischen Gebrauch bereitzustellen. Als umso glücklicher ist der Umstand zu werten, dass sich einige Zeit nach Beendigung der Aufnahmen in Amorbach, Meisenheim und Simmern zufällig die Möglichkeit ergab, in der Stadtkirche »Zur Gotteshilfe« im thüringischen Waltershausen die dortige VH zu analysieren. Es handelt sich hierbei zwar um kein Originalregister aus der Erbauungszeit der Orgel (TOBIAS HEINRICH GOTTFRIED TROST 1730/55), dafür jedoch um eine Kopie des *Registers* in Weingarten, die dort im Rahmen der Restaurierung 1995/98 eingebaut wurde.

Die VH der Gebrüder STUMM entsprechen weitestgehend der französischen Ursprungsausführung (Abb. 6.2, vgl. auch Abb. 2.1 auf S. 11): Ihr *Resonator* ist in seiner unteren Hälfte (bis knapp oberhalb der *Stiefel*-Mündung) umgekehrt konisch, oben zylindrisch und wird an seiner Oberseite von einem aufgelöteten flachen Metallblatt etwa zur Hälfte verschlossen, was eine Dämpfung bewirkt. Mit diesem kann auch die Tonhöhe fein abgestimmt werden. Die *Kehlen* sind ebenfalls nach französischem Vorbild

<sup>76</sup>Die Äußerung von CHRISTIAN F. D. SCHUBART über die klangliche Qualität der VH der Orgel der Mannheimer Konkordienkirche aus der Werkstatt Stumm (vgl. das Zitat auf S. 15) unterstützte die Entscheidung für Instrumente dieser Orgelbauerfamilie. Leider existieren diese Orgel und damit auch ihre beiden VH nicht mehr.

<sup>77</sup>Diese Orgel besaß ursprünglich wie auch die Orgel der Gebrüder Stumm in der Mannheimer Konkordienkirche gleich zwei VH-*Register* [Eppelsheim 1971, 47, Fußnote 18].



**Abbildung 6.2:** Ton  $g^0$  (ohne Zungenblatt) einer Vox humana der Orgelbauer JOH. PHILIPP & JOH. HEINRICH STUMM (Amorbach, 1782).

geformt. Wiewohl alle drei VH demselben Konstruktionsprinzip folgen, zeigen sie dennoch geringfügige Abweichungen zueinander. Dies ist dem Umstand geschuldet, dass im 18. Jahrhundert eine Fertigung mit solcher Präzision, wie wir es heute gewohnt sind, technisch schlicht unmöglich war.<sup>78</sup>

Das regional wohl bekannteste und auch theoretisch nächstgelegene Instrument der Gebrüder STUMM in der Ev. Paulskirche Kirchheimbolanden (1745) konnte nicht mit einbezogen werden, da seine VH mit sehr hoher Wahrscheinlichkeit nicht mehr dem Ursprungszustand entspricht [Bonkhoff 1984, 60]. Ähnliches gilt für die grundsätzlich sehr gut erhaltene STUMM-ORGEL in der Mainzer Augustinerkirche (1773), bei deren VH es sich jedoch um eine moderne Rekonstruktion handelt.

Obwohl alle drei ausgewählten Orgeln von denselben Orgelbauern stammen und sich die VH auch in allen drei Fällen im selben Werk der Orgel, dem *Positiv*, befindet, ist ihre Aufstellung nicht identisch. Dies hängt mit der Konstruktion der jeweiligen *Windlade* zusammen: Während es sich in Amorbach um eine Windlade mit *chromatischer Pfeifenaufstellung* handelt, sind die Pfeifen in Meisenheim und Simmern *diatonisch* aufgestellt (Tabelle 6.1). Die VH in Waltershausen war ebenfalls *diatonisch* aufgestellt.

Die Aufnahmen wurden nach folgenden Grundsätzen durchgeführt:

- Von jedem der zu vermessenden VH-Register werden über den gesamten Umfang der Klaviatur sämtliche Töne C und G (C, G,  $c^0$ ,  $g^0$ ,  $c^1$ ,  $g^1$ ,  $c^2$ ,  $g^2$ ,  $c^3$ ) einzeln als Haltetöne inklusive An- und Absprache aufgenommen (= 9 Einzelaufnahmen).

<sup>78</sup> Beispielsweise konnten Metalle seinerzeit nicht so rein wie heute verhüttet werden. Auch Metalllegierungen schwankten in ihrer Zusammensetzung zumeist um wenige Prozent. Beides hat ebenfalls einen gewissen Einfluss auf den konkreten Klang der jeweiligen Pfeifen.

**Tabelle 6.1:** Schema der Tonfolge einer Windlade (Umfang  $C-c^3$ ) mit chromatischer (oben) bzw. diatonischer Pfeifenaufstellung (unten).

C	Cis	D	Dis	E	F	Fis	G	Gas	A	B	H	c <sup>0</sup>	...	c <sup>2</sup>	cis <sup>2</sup>	d <sup>2</sup>	dis <sup>2</sup>	e <sup>2</sup>	f <sup>2</sup>	fis <sup>2</sup>	g <sup>2</sup>	gis <sup>2</sup>	a <sup>2</sup>	b <sup>2</sup>	h <sup>2</sup>	c <sup>3</sup>
C	D	E	Fis	Gas	B	c <sup>0</sup>	d <sup>0</sup>	e <sup>0</sup>	fis <sup>0</sup>	gis <sup>0</sup>	b <sup>0</sup>	...	c <sup>3</sup>	...	h <sup>0</sup>	a <sup>0</sup>	g <sup>0</sup>	f <sup>0</sup>	dis <sup>0</sup>	cis <sup>0</sup>	H	A	G	F	Dis	Cis

- Aufnahmedauer jeweils ca. 5 Sekunden. (*Zungenpfeifen* besitzen eine charakteristische und deutlich wahrnehmbare An- und Absprache-Phase, daher war es notwendig, durch eine solche verhältnismäßig lange Signaldauer eine ausreichend lange Periode des stabilen Klanges zu erreichen.)
- Vom Referenzton  $g^0$  werden die Maße des *Resonators* und – soweit möglich<sup>79</sup> – diejenigen der *Zunge* und der *Kehle* aufgenommen.

Wie bei jedem lingualen Orgelregister unterscheiden sich auch die VH-Pfeifen unterschiedlicher Tonhöhe durch längere *Zungenblätter* und *Kehlen* bei abnehmender Tonhöhe voneinander. Während bei »normalen« *Zungenregistern* jedoch auch die Länge des *Resonators* proportional zur Tonhöhe ansteigt, bleibt sie bei VH-Pfeifen relativ konstant (vgl. auch das folgende Zitat).

Da in der Fachliteratur die Tenorlage der VH ( $-c^0$ – $c^1$  bzw. -130–260 Hz) als der menschlichen Stimme am ähnlichsten beschrieben wird, wurde  $g^0$  (-200 Hz) als Referenzton gewählt [vgl. Locher 1887, 125; Frotscher 1927, 54 ].

In Simmern waren – anders als bei den übrigen aufgenommenen Instrumenten – auch die *Lingualregister Trompette* (TR) und *Cromorne* (CR) noch original erhalten. Diese unterscheiden sich konstruktiv und damit auch klanglich sowohl von der VH als auch untereinander, es handelt sich aber gleichwohl um Pfeifen mit demselben Funktionsprinzip wie die VH. Von ihnen wurden jeweils die Töne C und  $g^0$  als Referenz- und Vergleichsaufnahme mit in die Untersuchung einbezogen. Die Gebrüder STUMM leiteten ihre VH-Bauform aus derjenigen ihres *Cromorne* ab, wie es auch bei der französischen Ursprungsausführung üblich war:

<sup>79</sup>Zungenpfeifen – zumal historische – sind hochempfindlich gegenüber mechanischen Erschütterungen. Diese wirken sich unmittelbar auf den Klang aus. Zur fotografischen Aufnahme und Vermessung von *Kehle* und *Zungenblatt* muss jedoch die Zungenpfeife aus ihrer Halterung genommen und in ihre Einzelteile zerlegt werden, was unweigerlich solche Störungen hervorruft. Die Messungen konnten daher nur in Gegenwart eines Orgelbauers durchgeführt werden, der in der Lage war, die hierdurch entstandenen intonatorischen Mängel sogleich wieder zu beheben.

»Vox humana 8' ist ein Regal in der Form des Krummhorns [Cromorne] mit einem konischen Unter- und einem zylindrischen Oberteil. Als Regal hat es einen kürzeren Becher [Resonator], dessen zylindrischer Teil immer kürzer wird [mit ansteigender Tonhöhe], während der konische Teil annähernd konstant bleibt und dadurch in den höheren Lagen wesentlich größer ist.« [Bösken 1981, 33]

Jeder Raum beeinflusst durch seine jeweils vollkommen individuelle Größe, Form und Ausstattung und den Aufstellungsort der Orgel deren Klang erheblich. Jeder Orgelbauer passt die einzelnen *Register* einer Orgel stets individuell an den einzigartigen Raum an, in dem diese steht. Die klangliche Abstimmung einer VH ist also auf ihre Wirkung im Raum unter Berücksichtigung von dessen filternden Eigenschaften berechnet und nicht auf den Klangeindruck in direkter Nähe zur Pfeife.

Um die Messergebnisse vergleichbar zu gestalten, musste jedoch dieser akustische Einfluss des individuellen Raumes möglichst vermieden werden, wiewohl nicht ausgeschlossen werden kann, dass just diese »filternden« akustischen Parameter des Raumklanges mitverantwortlich für eine mehr oder weniger »authentische« Wirkung des *Registers* sind. Eine Aufnahme der Probetöne sowohl direkt an den Pfeifen als auch aus dem Raum heraus wäre daher wünschenswert und sinnvoll gewesen.

Aufgrund der beschriebenen Unterschiedlichkeit verschiedener Räume in allen relevanten Parametern wäre allerdings ein immenser technischer und zeitlicher Aufwand vonnöten gewesen, um auch nur annähernd vergleichbare Klangproben der Pfeifen vom Raum aus zu erhalten (exakte Bestimmung der absoluten Höhe des Mikrofons im Raum, seiner relativen Höhe zur Schallaustrittsfläche der Orgel, Abstand von dieser, Abstand zu schallreflektierenden Wänden und Gewölben, Wissen über deren exakten materiellen Aufbau etc.). Daher wurde auf diese Aufnahmen verzichtet.

Während der Aufnahmen zeigten sich mehrere zuvor nicht mit einkalkulierte Faktoren: Die *Windladen* einer Orgel, auf denen die Pfeifen der Breite nach stehen, sind mit Mäßen von ca. 3 Metern Breite in aller Regel zu ausgedehnt, als dass Pfeifen von einem Mikrofonstandort aus angemessen erfasst werden könnten.<sup>80</sup> Von der daraus resultierenden ursprünglichen Intention, die Aufstellung des Mikrofons für jede Pfeife individuell anzupassen, musste jedoch vor Ort aufgrund der jeweils sehr beengten und damit höchst unfallträchtigen Umgebung abgesehen werden. Stattdessen wurde das Mikrofon in seiner Position nur je zweimal verändert: Die Pfei-

<sup>80</sup>Bei einer Aufnahme aus dem Kirchenraum heraus nivelliert sich diese Breitenverteilung.

fen, die auf der linken bzw. rechten Hälfte der Windlade standen, wurden mit einer jeweils auf diesen Windladenteil gemittelten Mikroaufstellung aufgenommen.

Im Fall der STUMM-ORGEL in Amorbach konnten daher die Töne C-c<sup>1</sup> und g<sup>1</sup>-c<sup>3</sup> mit je einer Mikrofonpositionierung aufgenommen werden, bei den übrigen jeweils alle Töne C und G (vgl. Tab. 6.1). Die Aufnahmen wurden angefertigt mit einem Marantz solid state PMD670-Aufnahmegerät (Netzbetrieb) und einem Beyer Dynamic MCE 86N(C)-Richtmikrofon. Die Aufzeichnung erfolgte auf einer 2 GB Kingston Elite Pro-Speicherkarte. Das Mikrofon wurde jeweils in ca. 0,5 m Höhe senkrecht über den Resonatoren der Pfeifen installiert.

## 6.2.2 Analyse

Die Analyse der Aufnahmen geschah mit der freien Analysesoftware praat 5.3.19 unter Zuhilfenahme eines Skripts zur Formantermittlung von Prof. Dr. BEAT SIEBENHAAR (Leipzig). Im Laufe der Auswertung stellte sich heraus, dass die Orgeln in Meisenheim und Simmern, anders als die in Amorbach, noch in ihrer originalen Tonhöhe stehen, was selbst bei weitgehend erhaltenen Orgeln der Barockzeit nicht der Regelfall ist. Während in Amorbach die Stimmtonghöhe der modernen von a<sup>1</sup> = 440 Hz entspricht, sind die anderen beiden STUMM-Orgeln auf etwa a<sup>1</sup> = 462 Hz eingestimmt, stehen also um einen knappen Halbton höher als die Orgel in Amorbach. Dieser in Meisenheim und Simmern angewendete Stimmtong war im Barock eine übliche Stimmtonghöhe für kleinere Orgeln [Busch 2007, 742].<sup>81</sup> Die Amorbacher Orgel war ursprünglich sogar auf den erheblich tieferen *Chorton* (bzw. *Cornetton*, siehe Fußnote 81) eingestimmt, klang also noch um etwa einen Halbton tiefer als heute und damit einen Ganzton tiefer als die Orgeln in Meisenheim und Simmern (pers. Mitteilung von DR. HANS-WOLFGANG THEOBALD).<sup>82</sup> Auch die Orgel in Waltershausen steht noch in

---

<sup>81</sup>Die Gebr. Stumm bezeichneten diesen Stimmtong als *Cornetton*, wohingegen er sonst zu dieser Zeit als *Chorton* bekannt war, während unter *Cornetton* eine andere, erheblich tiefere Stimmtonghöhe verstanden wurde, die von den STUMM wiederum als *Chorton* bezeichnet wurde. Die STUMM bedienten sich in dieser Vertauschung der Begriffe einer schon zu ihrer Zeit altertümlichen Nomenklatur, die im Widerspruch zur zeitgenössischen Konnotation dieser Bezeichnungen zu bestimmten Stimmtonghöhen stand [Bösken 1981, 110].

<sup>82</sup>Für die Höherstimmung auf die heutigen Werte war im Falle der uns interessierenden *Zungenregister* kein Austausch nötig. Lingualpfeifen können in gewissem Rahmen umgestimmt werden.



**Tabelle 6.2:** *Tonhöhen der einzelnen VH-Töne (in Hz).*

Ton	Amorbach	Meisenheim	Simmern	Waltershausen
C	66	70	69	69
G	98	105	102	103
c <sup>0</sup>	132	141	136	139
g <sup>0</sup>	197	210	205	208
c <sup>1</sup>	263	281	274	277
g <sup>1</sup>	395	411	408	415
c <sup>2</sup>	527	562	548	554
g <sup>2</sup>	790	844	818	832
c <sup>3</sup>	1054	1124	1093	1108

ihrer historischen hohen Stimmung.<sup>83</sup> Tab. 6.2 zeigt synoptisch die Werte der aktuellen Stimmtöne für alle vier aufgenommenen VH-Töne.

LOTTERMOSER bringt in seiner Untersuchung der VH in Weingarten auch eine spektrale Analyse von deren Ton C, erklärt jedoch zugleich, dass als Regelfall ein Energiemaximum auf dem 7. Teilton, und nicht wie in Weingarten, auf dem 6. zu erwarten sei [Lottermoser 1983a, 135] (vgl. 6.3 auf Seite 137). An anderer Stelle nennt er zudem Formantwerte von -700 Hz, -1800 Hz und -3000 Hz für den Ton C der Weingartener VH [Lottermoser 1983a, 125 f.] (vgl. auch Tab. 6.3).

Ein Vergleich dieser Werte mit den für die vier VH ermittelten Werten ist schwierig, da LOTTERMOSER seinerzeit nur sehr unzureichende Messmethoden zur Verfügung standen. Daher sind die von ihm mitgeteilten Werte nur näherungsweise zu interpretieren, wie auch bereits die gerundeten Werte erkennen lassen. Eine Synopse der Formanten F<sub>1</sub> bis F<sub>3</sub> für die Werte aus Weingarten und die vom Autor vermessenen VH (Tab. 6.3) zeigt, dass die Werte für Waltershausen erwartungsgemäß abweichend ausfallen, was durch die zusätzliche Benutzung des *Registers Gedeckt 8'* zu erklären ist. Die drei STUMM-Register weisen untereinander gewisse Abweichungen auf, kommen für F<sub>2</sub> und F<sub>3</sub> aber den schematischen Werten für Weingarten nahe. Lediglich für F<sub>1</sub> besteht eine erhebliche Differenz.

Vergleicht man die Energiemaxima aus Amorbach, Meisenheim und Simmern mit dem für Weingarten mitgeteilten, so findet sich das von LOT-

<sup>83</sup>Die geringfügigen Differenzen der Tonhöhen zwischen den drei in hoher Stimmung stehenden Orgeln resultieren aus den jeweils angewendeten geringfügig unterschiedlichen *Temperierungen*. Die Orgel in Amorbach besitzt als einzige der vier untersuchten eine moderne *gleichschwebende Temperatur*.

**Tabelle 6.3:** *Formantwerte für den Ton C der VH in Amorbach, Meisenheim, Simmern, Waltershausen und Weingarten.*

	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>
Amorbach	1036	1628	2856
Meisenheim	1012	1548	2550
Simmern	983	1637	2823
Waltershausen	453	1587	2573
Weingarten	700	1800	3000

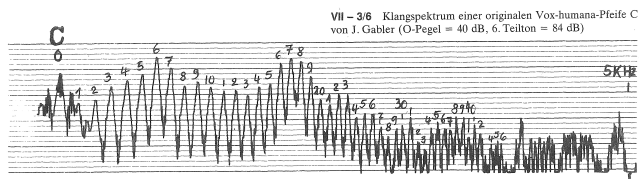
**Tabelle 6.4:** *Teiltöne mit absolutem Energiemaximum*

	C	G	c <sup>0</sup>	g <sup>0</sup>	c <sup>1</sup>	g <sup>1</sup>	c <sup>2</sup>	g <sup>2</sup>	c <sup>3</sup>
Amorbach	15	13	4	3	3	9	5	3	1
Meisenheim	6	5	16	9	7	5	4	3	1
Simmern	5	5	10	8	7	5	2	1	2
Simmern CR	15			7					
Simmern TR	9			1					

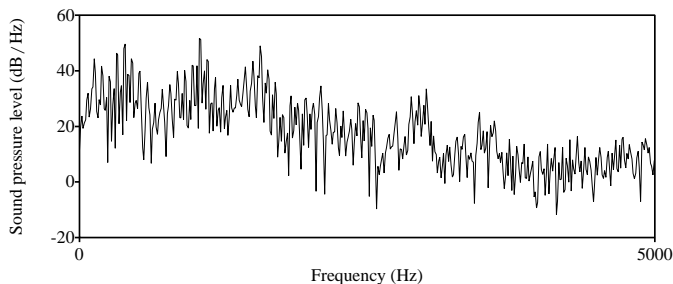
TERMOSER beschriebene Energiemaximum auf dem 7. Teilton bei lediglich zwei Einzeltönen, jeweils c<sup>1</sup> aus Meisenheim und Simmern. Sämtliche übrigen Töne der drei STUMMSCHEN VH weisen andere, scheinbar mehr oder minder arbiträr verteilte Energiemaxima auf (Tab. 6.4 und Abb. 6.6 – 6.8). Abb. 6.6–6.8 zeigen auch, dass der beobachtbare Unterschied zwischen der VH in Waltershausen und den drei STUMM-*Registern* nicht einzig auf die zusätzliche Verwendung des *Gedeckt 8'* zurückzuführen ist. Vergleicht man beispielsweise die Spektra des Tons C aus Waltershausen und Simmern (einmal ohne und einmal mit *Gedeckt 8'* aufgenommen), so erweist sich, dass sich diese beiden *Register* auch dann noch spektral voneinander deutlich unterscheiden, wenn sie beide mit *Gedeckt 8'* gespielt werden.

Abb. 6.9 bietet ergänzend eine synoptische Darstellung der Teiltonintensitäten für die in den vorangehenden Spektren dargestellten Töne C der vier VH. Hierzu wurde jeweils für die mittlere stabile Phase der Töne mithilfe von *praat* (Version 5.3.39) ein *Long term average spectrum* (LTAS) berechnet. Die vier LTAS wurden anschließend mithilfe der Zeichenfunktion von *praat* quasi dreidimensional als übereinandergelegte Säulendiagramme dargestellt und farblich codiert (pro VH eine Farbe): Für jeden der 50 Messpunkte zwischen 0 und 5 kHz (jeweils eine Säule) wird diejenige VH,

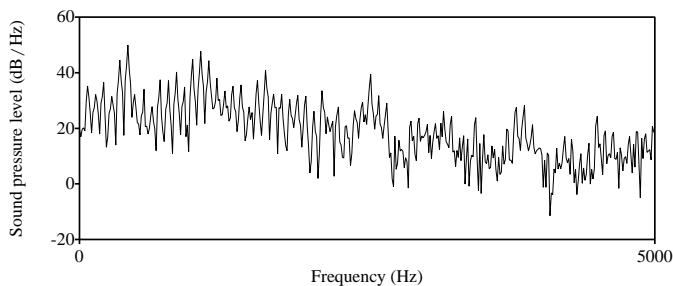
die bei diesem Teilton die geringste Intensität aufweist, im »Vordergrund« dargestellt, diejenige mit der höchsten Intensität im »Hintergrund«.



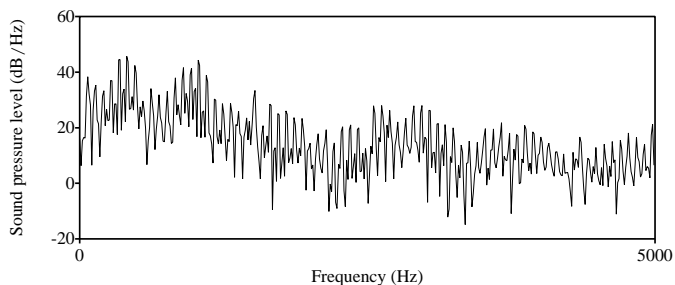
**Abbildung 6.3:** Spektrum von 0 bis 5 kHz für den Ton C der VH in Weingarten [Lottermoser 1983a, 135].



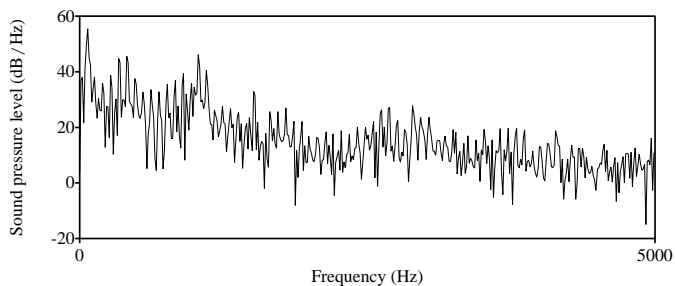
**Abbildung 6.4:** Amorbach, VH solo, Ton C, Spektrum über 4 Perioden.



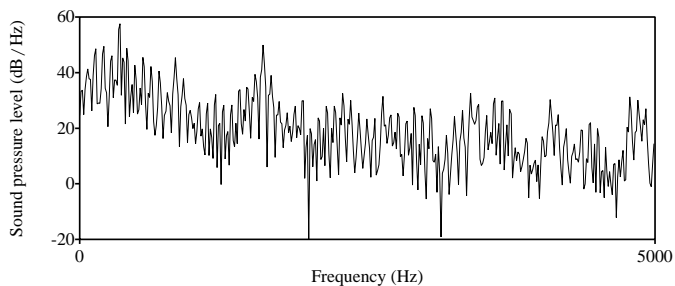
**Abbildung 6.5:** Meisenheim, VH solo, Ton C, Spektrum über 4 Perioden.



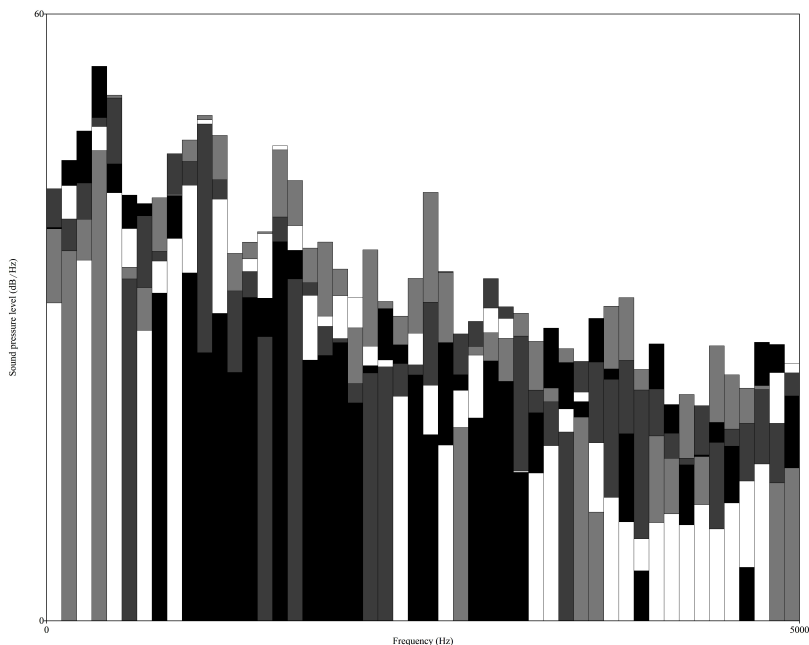
**Abbildung 6.6:** *Simmern, VH solo, Ton C, Spektrum über 4 Perioden.*



**Abbildung 6.7:** *Simmern, VH + Gedeckt 8', Ton C, Spektrum über 4 Perioden.*



**Abbildung 6.8:** *Waltershausen, VH + Gedeckt 8', Ton C, Spektrum über 4 Perioden.*

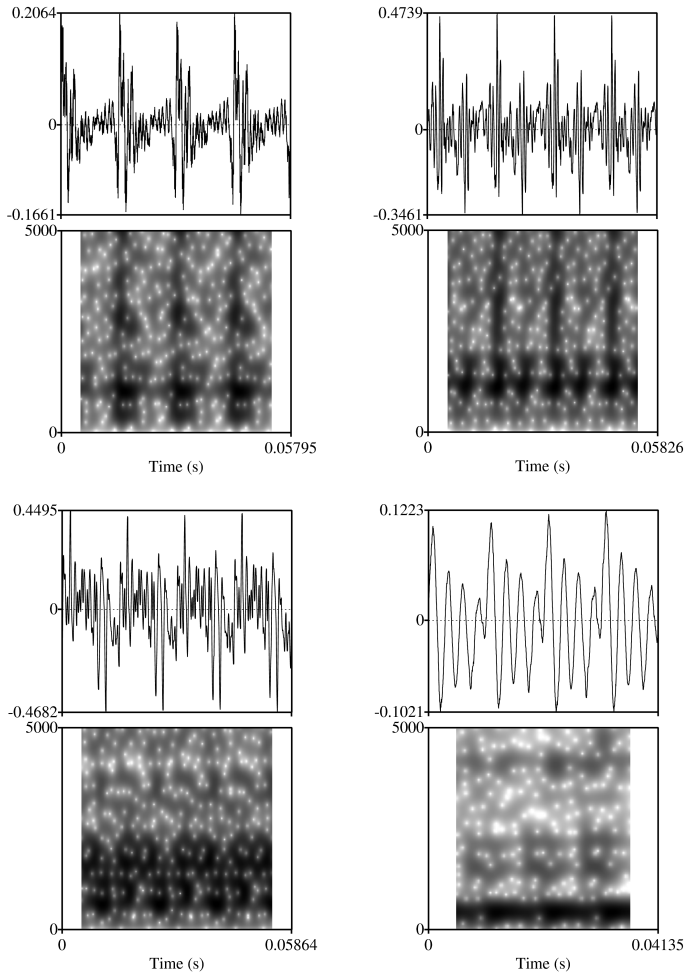


**Abbildung 6.9:** *Überinandergelegte Long Term Average Spectra (LTAS) von 0 bis 5 kHz des Tones C der VH aus Amorbach (weiß), Meisenheim (hellgrau), Simmern (dunkelgrau) und Waltershausen (schwarz).*

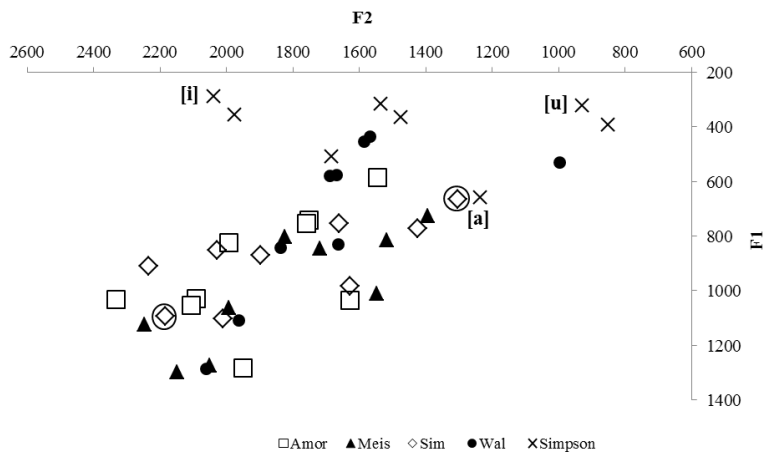
Der Befund, dass die ermittelten Formant-Maxima für Amorbach stark von denen der anderen beiden Stumm-Organen abweichen, könnte der Tatsache geschuldet sein, dass dieses Register verhältnismäßig häufige und anscheinend tiefgreifende Veränderungen erfahren hat. Diese wurden bei der letzten Restaurierung zwar weitgehend rückgängig gemacht, anscheinend blieben aber trotzdem akustische Folgen der Umbauten erhalten.

Ein Vergleich der Oszillogramme und Spektrogramme des Tons C der Register VH, *Trompette* und *Cromorne* aus Simmern mit denjenigen eines menschlichen Sprechers (Abb. 6.10) zeigt jedoch, dass die Energieverteilung bei der VH derjenigen der menschlichen Stimme noch am nächsten kommt. Der Ton C der Simmerner VH war in den nachfolgend geschilderten Perceptionsexperimenten sehr häufig mit einer /ø:/-ähnlichen Vokalqualität assoziiert worden. Daher wurde als Vergleichsbeleg der Nukleus-Vokal des deutschen Wortes <Köder> gewählt, gesprochen von einem deutschen Muttersprachler mit sehr tiefer Stimme. Diese war im Vokal

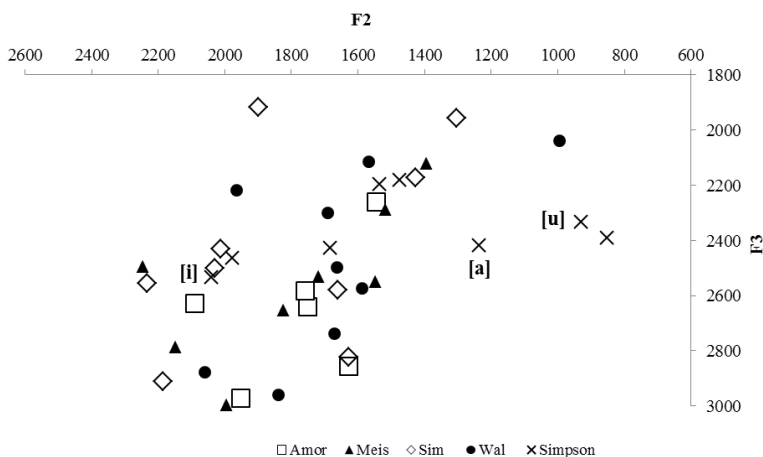
mit 97 Hz jedoch nach wie vor erheblich höher als die Frequenz des Tons C in Simmern mit 69 Hz. Für den Ton G, mit 102 Hz praktisch gleichhoch mit der Stimme des Vergleichssprechers, lagen indes keine Aufnahmen der *Register Trompette* und *Cromorne* vor.



**Abbildung 6.10:** Oszillo- und Spektrogramm-Ausschnitte des Tons C (69 Hz) der Register VH (links oben), Cromorne (rechts oben) und Trompette (links unten) aus Simmern sowie für den Vokal /ø:/, gesprochen von einem männlichen deutschen Muttersprachler, im Wort <Köder> (97 Hz) (rechts unten).



**Abbildung 6.11:** Werte für  $F_1$  und  $F_2$  aller aufgenommenen VH-Töne sowie für die Referenzwerte für gesprochene Vokale eines männlichen Sprechers des Deutschen [nach Simpson 1998]. Die umkreisten Datenpunkte bezeichnen die in den im Folgenden beschriebenen Perceptionsexperimenten besonders konsistent bewerteten Stimuli, korrespondierend zu Tab. 6.10 und 6.12.



**Abbildung 6.12:** Werte für  $F_2$  und  $F_3$  aller aufgenommenen VH-Töne sowie für die Referenzwerte eines männlichen Sprechers des Deutschen [nach Simpson 1998].

**Tabelle 6.5:** Formantwerte (Hz)  $F_1$  und  $F_2$  für den VH-Ton  $g^0$  in Meisenheim in verschiedenen Registrierungen.

	$F_1$	$F_2$	$F_1/F_2$
VH solo	862	1846	2,14
VH + Gedeckt 8'	670	1847	2,80
VH + Gedeckt 8' + Tremulant	742	1847	2,49

Alle aufgenommenen VH-Töne weisen ausgeprägte und charakteristische Formantstrukturen auf. Dies ist jedoch, wie Abb. 6.10 zeigt, kein Alleinstellungsmerkmal der VH. Auch andere *Zungenregister* weisen formantartige Strukturen auf, die allerdings in ihrer Energieverteilung einer menschlichen Stimme weniger ähneln.

Bemerkenswert ist, dass sich im Spektrogramm aller Töne der drei STUMM-Register keinerlei Grundfrequenz erkennen und auch nicht messen lässt.<sup>84</sup> Dies unterstützt den Eindruck aus den obigen Spektren, bei denen jeweils der 1. Teilton nicht übermäßig stark ausgeprägt ist. Die Messung der  $F_0$  war allerdings bei den Aufnahmen, die vom Ton  $g^0$  mit VH + Gedeckt 8' gemacht wurden, möglich. Der hier für Meisenheim gemessene Wert von  $g^0=210$  Hz korreliert vollkommen mit dem Stimmton von  $a^1=462$  Hz.

Einen Erklärungsansatz für den nahezu maximal gedämpften 1. Teilton gibt ELLERHORST: Die *Resonatoren* filtern je nach Konstruktion verschiedene Teiltöne verschieden stark [Ellerhorst 1936, 322]. Diejenigen von *Regalen* scheinen gerade den Grundton besonders effektiv zu filtern. Die konkreten Formantwerte sind jedoch zunächst wenig aussagekräftig (vgl. Tabelle 6.3). Insbesondere die Werte für  $F_1$  finden keinerlei Entsprechungen bei deutschen Vokalen (vgl. [Simpson 1998, 199 f.]). Auch die Berechnung eines Quotienten ( $F_1/F_2$ ) erbrachte keine proportionalen Werte, die denen deutscher Vokale vergleichbar sind. Trägt man die VH-Formantwerte jedoch grafisch gegen diejenigen der deutschen gespannten Langvokale eines menschlichen Sprechers auf, so ergibt sich ein etwas klareres Bild (Abb. 6.11 und 6.12):

Die Werte für den menschlichen Sprecher heben sich im Allgemeinen deutlich von den VH-Werten ab. Die Töne aus Waltershausen reflektieren verstärkt Veränderungen im  $F_1$ -Bereich, wohingegen die drei STUMM-Register eher Variation im Bereich von  $F_2$  aufweisen. Gemeinsam ist jedoch

<sup>84</sup>Die in Tab. 6.2 mitgeteilten Grundfrequenzen wurden über das genannte praat-Skript ermittelt.



allen vier *VH-Registern*, dass sie sich diagonal von den Referenzwerten fortbewegen. Hinsichtlich der  $F_3$  weisen alle VH eine breitere Streuung auf als es beim menschlichen Sprecher der Fall ist.

Die zusammen mit *Gedeckt 8'* in Meisenheim gemessenen absoluten Werte liegen in der Nähe von  $/\varepsilon/$ , der Quotient jedoch ist ein deutlich anderer (2,80 statt 3,20). Unter Zuziehung des *Tremulanten* veränderte sich ausschließlich  $F_1 - F_0$  und  $F_2$  blieben hingegen konstant (Tab. 6.5).

Auch bei der Ermittlung der Formantwerte für die Töne C-c<sup>1</sup> differieren die Formantwerte für die einzelnen Töne zwischen Meisenheim und Amorbach teilweise erheblich. Eine händische Nachmessung erbrachte nur graduelle Unterschiede zu den via Skript gemessenen Werten. Hinzu kommt, dass im Falle Amorbach noch nicht einmal die mehrfach ermittelten Werte ein und desselben Tones übereinstimmen: Der Ton c<sup>1</sup> war in Amorbach zu Kontrollzwecken zweimal aufgenommen worden, da er durch die dortige *chromatische Pfeifenstellung* auf der *Windlade* genau an der Grenze der mit einer Mikrofonpositionierung aufgenommenen Tonbereiche lag.<sup>85</sup> Während die  $F_0$  mit 263 Hz praktisch konstant blieb, differieren die Werte für  $F_1$  erheblich, die für  $F_2$  und  $F_3$  noch geringfügig.

## 6.3 Perzeptionsexperimente

### 6.3.1 Fragestellungen und Methoden

Die im vorangegangenen Abschnitt vorgestellten Daten lassen nicht ohne Weiteres erkennen, warum der Klang von *VH-Registern* ganz offenbar lange Zeit nicht nur als menschenähnlich postuliert, sondern zumindest vereinzelt auch so empfunden wurde.<sup>86</sup> Es bliebe also die Frage, ob auch heutige Hörer den Klang der konstruktiv allem Anschein nach stark der historischen französischen Ursprungsversion verhafteten *VH-Register* der Gebrüder STUMM bzw. dem Nachbau der betont »anatomie-nah« konstruierten GABLER-VH aus Waltershausen als stimmähnlich empfinden. Aufgrund der nahezu vollständigen Periodizität des VH-Klangs liegt hierbei ein Vergleich mit Vokalen natürlich besonders nahe.

#### 6.3.1.1 Experiment Nr. 1

Es wurden drei aufeinander aufbauende Perzeptionstests konzipiert. Im ersten, als Pilottest anzusehenden, wurden insgesamt 40 Stimuli verwen-

<sup>85</sup>vgl. Tab. 6.1 auf S. 132

<sup>86</sup>Unzweifelhaft kann es hierbei nicht um einen Eindruck von gesprochener Sprache gehen, sondern einzig um den von Gesang.

det. Diese setzten sich aus 36 VH-Stimuli<sup>87</sup> und vier Kontrollstimuli der bereits erwähnten ebenfalls originalen *Lingualregister Trompette* und *Cromorne* aus Simmern wie folgt zusammen:

- *Vox humana* Amorbach (9 Stimuli)
- *Vox humana* Meisenheim (9 Stimuli)
- *Vox humana* Simmern (9 Stimuli)
- *Vox humana* Waltershausen (9 Stimuli)
- *Trompette* Simmern (2 Stimuli)
- *Cromorne* Simmern (2 Stimuli)

Diese Stimuli wurden 20 Versuchspersonen, alle Germanisten am »Institut für Deutsche Sprache« (Mannheim), in randomisierter Reihenfolge via Kopfhörer vorgespielt. Jeder Stimulus wurde im Verlauf des Experiments nur einmal präsentiert, konnte aber beliebig oft angehört werden. Die Versuchspersonen sollten entscheiden, ob der jeweils gehörte Stimulus einem Vokal ähnele und diesen dann ggf. dem Versuchsleiter mündlich benennen. Die Versuchspersonen wurden gebeten, die Vokalqualität nach Möglichkeit in den Kategorien des *Internationalen phonetischen Alphabets* (IPA) anzugeben. Das Antwortformat »kein Vokal« war explizit ebenfalls möglich. Die Experimentdauer variierte aufgrund der unterschiedlichen »Entscheidungsfreudigkeit« der Teilnehmer stark zwischen ca. 15 Minuten und etwa 45 Minuten.

### 6.3.1.2 Experiment Nr. 2

Das zweite Experiment baute auf dem ersten auf, trug aber den Erfahrungen aus diesem Pilotversuch Rechnung. Um eine möglichst große Anzahl von Versuchspersonen zu gewinnen, wurde es als Web-Experiment konzipiert [Draxler 2011].<sup>88</sup> Es konnten schließlich 29 deutsche Muttersprachler, sowohl Linguisten als auch Laien, zur Teilnahme gewonnen werden, von denen niemand am ersten Experiment teilgenommen hatte.

Im nun modifizierten Experimentdesign wurden lediglich die Stimuli aus Simmern und Waltershausen (2x9+4=22) verwendet, da diese im Pilotexperiment die kontrastivste Bewertung erfahren hatten. Die Reduktion

<sup>87</sup> 4x9 Stimuli = Alle Einzeltöne C, G, c<sup>0</sup>, g<sup>0</sup>, c<sup>1</sup>, g<sup>1</sup>, c<sup>2</sup>, g<sup>2</sup>, c<sup>3</sup> für alle vier VH-Register.

<sup>88</sup> Mein Dank gilt Dr. CHRISTOPH DRAXLER (Ludwig-Maximilians-Universität, München) für seine große Hilfsbereitschaft.

der Stimulus-Anzahl erfolgte auch, um trotz zweier Wiederholungen der Stimuli innerhalb des Experiments dessen Dauer auf eine knappe Viertelstunde begrenzen zu können. Im Laufe des Pilotexperiments hatte sich gezeigt, dass eine deutlich längere Experimentdauer nicht mit der Konzentrationsfähigkeit einiger Teilnehmer in Einklang zu bringen war.

Die Stimuli wurden auf eine vokaltypische Länge von 400 ms aus der Mitte heraus gekürzt und wiederum in randomisierter Anordnung wiedergegeben. Jeder Stimulus wurde im Verlauf des Experiments drei Mal zur Bewertung gestellt. Die Versuchspersonen wurden aufgefordert, die Stimuli via Kopfhörer oder externe Lautsprecher in ruhiger Umgebung anzuhören und jeweils einer der Vokalkategorien I, Ü, E, Ä, Ö, A, O, U zuzuordnen. Von der im Pilotversuch angebotenen Möglichkeit, in IPA-Kategorien zu antworten, hatten nur drei Versuchspersonen effektiv Gebrauch gemacht, es wurde daher nun nur noch auf orthografische Kategorien zurückgegriffen. Das Antwortformat »kein Vokal« wurde nicht mehr zugelassen. Zu Beginn des Experiments wurden sechs ebenfalls auf 400 ms gekürzte Stimuli aus Meisenheim und Amorbach als Training vorgeschaltet. Diese wurden nicht mit in die Auswertung einbezogen.

Durch einen Fehler in der Datenübermittlung wurde im zweiten Experiment der Stimulus c<sup>0</sup> aus Waltershausen nicht in allen Fällen korrekt abgespielt und konnte somit nicht in die Auswertung miteinbezogen werden.

### 6.3.1.3 Experiment Nr. 3

Das dritte Experiment wurde mit einigem zeitlichen Abstand zu den vorangehenden beiden durchgeführt und berücksichtigte die aus diesen gewonnenen Erkenntnisse und neuen Hypothesen. Es sollte nun explizit überprüft werden, inwieweit die Vokallasoziiierung der VH-Töne nicht allein in den gemessenen Formanten F<sub>1</sub>–F<sub>3</sub> begründet liegt, sondern die jeweilige Grundfrequenz einen merklichen Anteil hieran hat. Hierzu sollten die Grundfrequenzen und die Formanten der VH-Stimuli unabhängig voneinander untersucht werden. Dazu wurden die Stimuli aus dem zweiten Perzeptionstest durch klatt-synthetisierte Pendants ergänzt. Da sich in den vorangehenden Versuchen bereits gezeigt hatte, dass die Stimuli aus Simmern am konsistentesten von allen vier VH mit bestimmten Vokalqualitäten assoziiert werden konnten, wurden ausschließlich diese für das dritte Experiment verwendet. Diese Auswahl wurde auch getroffen, um die Anzahl der Stimuli und damit wiederum die Experimentdauer in einem vertretbaren Rahmen zu halten. Demnach fanden insgesamt 42 Stimuli aus

fünf Kategorien Verwendung.<sup>89</sup>

- I. VH-Stimuli aus Simmern (9 Stimuli).
- II. Klatt-synthetisierte Langvokale.  $F_0=100$  Hz,  $F_1$ ,  $F_2$  und  $F_3$  nach [Simpson 1998]. Es wurden nur die Vokalqualitäten verwendet, die als Assoziationskategorie im Perzeptionsexperiment 2 zugelassen worden waren (8 Stimuli<sup>90</sup>).
- III. Klatt-synthetisierte Kombinationen aus VH-Tönen und Vokalen:
  - a.  $F_0$  und  $F_1$ – $F_3$  nach I. (8 Stimuli<sup>91</sup>)
  - b.  $F_0 = 100$  Hz,  $F_1$ – $F_3$  nach I. (9 Stimuli)
  - c.  $F_0$  nach I.,  $F_1$ – $F_3$  nach [Simpson 1998], jeweils für die Vokalqualität, die im Perzeptionsexperiment Nr. 2 mit der jeweiligen Grundfrequenz am häufigsten assoziiert wurde (8 Stimuli).

Die synthetischen Stimuli wurden mithilfe zweier Software-Tools zur einfachen Klatt-Vokalsynthese erzeugt: Für die Kategorien II, IIIa und IIIc wurde das Webinterface des Speech Research Lab der University of Delaware verwendet.<sup>92</sup> Da dieser Syntheservice unvermittelt und dauerhaft eingestellt wurde, mussten die Stimuli für Kategorie IIIb mithilfe der Freeware IMSKPE (ims klatt parameter editor)<sup>93</sup>, Version 1.0.2, erzeugt werden.

Die der Experimentkonzeption zugrundeliegende Überlegung war, nur die gemeinhin als wesentlich zur Vokaldistinktion beschriebenen Parameter der VH-Töne, also Grundfrequenz und die Formanten  $F_1$ ,  $F_2$  und  $F_3$ , zu verwenden [vgl. Ashby & Maidment 2007, 71] und die der Tonerzeugung geschuldeten mechanischen Nebengeräusche zu vermeiden. Die Hypothese besagte, dass diese Stimuli der Kategorie III deutlich schlechter bewertet werden würden als die synthetisierten Vokale (Kategorie II), aber auch inkonsistenter als die originalen VH-Töne (Kategorie I).

Das Experiment wurde analog zum zweiten Perzeptionsexperiment konzipiert. Alle Stimuli wurden wieder drei Mal in randomisierter Reihenfolge präsentiert und konnten jedes Mal bis zu drei Mal angehört werden. Auch die Instruktionstexte, wählbare Vokalkategorien sowie die sechs initialen Trainingsstimuli waren mit denen des zweiten Experiments identisch. Es

<sup>89</sup>Zur besseren Unterscheidbarkeit werden im Folgenden die drei Experimente mit arabischen, die drei Kategorien des 3. Experiments mit römischen Ziffern bezeichnet.

<sup>90</sup>DerVokal /y/ wurde bei allen Stimuli aus Simmern nur in sehr geringem Umfang »identifiziert« und daher nicht verwendet.

<sup>91</sup>Der Ton  $c^3$  konnte aufgrund seiner hohen Grundfrequenz vom 1093 Hz mit der verwendeten Softwares nicht resynthetisiert werden.

<sup>92</sup><http://www.asel.udel.edu/speech/tutorials/synthesis/vowels.html>

<sup>93</sup><http://sourceforge.net/projects/imskpe/>

nahmen insgesamt 46 deutsche Muttersprachler am Experiment teil, von denen sich niemand bei einem der vorangehenden Experimente zur Verfügung gestellt hatte. Jedoch beendeten nur 35 Personen das Experiment. Acht Teilnehmer brachen das Experiment bereits binnen Sekunden (während der sechs Trainings-Stimuli) ab, zwei weitere nach etwa einer Minute. Die elfte Person brach nach ca. drei Minuten und 67 Inputs ab, was der Hälfte der Einzelstimuli entspricht. Von den 35 auswertbaren Teilnehmern waren sieben Phonetiker, neunzehn Linguisten und neun bezeichneten sich keiner der beiden Kategorien als zugehörig.

### 6.3.2 Ergebnisse

Über Teilergebnisse der in diesem Kapitel beschriebenen Versuchsreihe wurde bereits in [Brackhane & Trouvain 2013a] sowie [Brackhane & Trouvain 2013c] berichtet. Eine umfassende und ausführliche Darstellung kann jedoch erst in dem hier gegebenen Rahmen erfolgen.

Da es sich bei den vorgestellten Perzeptionsexperimenten nicht um voneinander unabhängige, sondern konzeptuell aufeinander aufbauende Untersuchungen handelt, werden die daraus gewonnenen Ergebnisse nicht getrennt, sondern synoptisch vorgestellt und im Folgenden diskutiert. Die Ergebnisse wurden in einer Serie von Kreuztabellen illustrierend zusammengefasst. Diese finden sich zur besseren Vergleichbarkeit in synoptischer Anordnung am Ende dieses Abschnittes ab S. 150.

Im ersten Perzeptionsexperiment wurde die angebotene Kategorisierung nach IPA nur sehr unregelmäßig genutzt. Die Antworten erfolgten stattdessen zumeist in orthografischen Klassen. Daher wurden die entsprechenden Vokale – phonetisch unterschiedlich, in der Orthografie jedoch durch denselben Buchstaben repräsentiert – in den jeweiligen Kreuztabellen zusammengefasst.

Alle drei Perzeptionstests weisen auf einen deutlichen Zusammenhang zwischen der Grundfrequenz eines Stimulus und der jeweils schwerpunktmäßig damit assoziierten Vokalqualität hin (Tab. 6.6–6.16). Auf eine Faustformel gebracht bedeutet dies: Je höher der dargebotene Ton war, desto sicherer wurde er als I (bzw. /i/ im Test Nr. 1) bezeichnet, je tiefer der Ton, desto eher wurde er als O benannt. Hierbei zeigt sich für die hinsichtlich der Grundfrequenz peripheren Stimuli eine größere Bewertungskonsistenz: Der Ton  $c^3$  aus Simmern beispielsweise wurde in den Experimenten in 80%, 84% bzw. 86% als I charakterisiert, während der nächsttiefere Ton  $g^2$  bereits zwischen den Qualitäten I und E schwankt, wenn auch mit starker Tendenz zum I. Der in der Mittelregion angesiedelte Ton  $c^1$  aus Simmern wurde in den Perzeptionsexperimenten Nr. 1+2 recht gleichmäßig als E, Ä,

Ö und A bezeichnet.<sup>94</sup> Im dritten Perzeptionsexperiment fiel das Votum zugunsten von Ä aus, wenn auch die im zweiten Experiment zu beobachtende Streuung grundsätzlich weiterhin vorhanden ist. Der äquivalente Ton  $c^1$  aus Waltershausen schwankt in seiner Bewertung zwischen /a,ɑ/ (Test 1) und U (Test 2).

Die Vergleichsstimuli der *Register Trompette* und *Cromorne* zeigen über alle Tonhöhen hinweg eine deutlich geringere Bewertungskonsistenz. Dies gilt insbesondere für den Ton  $g^0$ , der ja als Referenzton auf die menschliche Baritonstimme gewählt wurde.

In den beiden ersten Experimenten zeigten die Stimuli aus Simmern die deutlich größere Konsistenz hinsichtlich der Korrespondenz zwischen Ton und Vokalqualität. Insgesamt konnte der zweite Perzeptionstest die Ergebnisse des ersten erhärten, wobei die Bewertungen für einige Stimuli stabiler wurden, beispielsweise bei den Simmerner Stimuli  $c^0$  und  $g^1$ .

Im dritten Experiment veränderten sich die Bewertungs-Schwerpunkte mit  $g^1$  und  $c^2$  nur für zwei der neun VH-Stimuli aus Simmern merklich. In fünf Fällen – die vier tiefsten Stimuli und der höchste Stimulus – werden die Ergebnisse aus den vorangegangenen Experimenten bestätigt; bei den zuvor eher divers kategorisierten Tönen aus der Mittellage  $c^1$  und  $g^2$  wurden die Bewertungen sogar stabiler (Tab. 6.10 und 6.12). Ergänzend zu dieser Bestätigung der Ergebnisse der ersten Experimente sorgen die Bewertungen der synthetisierten Stimuli zudem für eine deutliche Evidenz der These eines Zusammenhangs zwischen Grundfrequenz und Vokalqualität.

Ein Blick auf die Bewertungen der Stimuli der Kategorie IIIc bestätigt die Vermutung, dass eine hohe Grundfrequenz einen deutlichen Anteil an der Identifizierung eines synthetischen Stimulus mit einer bestimmten Vokalqualität hat (Tab. 6.16): Alle vier Stimuli mit einer Grundfrequenz  $> 200$  Hz wurden als geschlossene Vorderzungenvokale identifiziert. Der Einfluss von  $F_0$  wird besonders bei den vier Stimuli deutlich, die sämtlich die Formantwerte des /ø/ erhielten bei unterschiedlichen Grundfrequenzen. Während die drei Stimuli mit einer  $F_0 \leq 200$  Hz mit abnehmender Frequenz zunehmend sicher als Ö identifiziert wurden, wurde der vierte /ø/-Stimulus mit einer  $F_0$  von 274 Hz mit großer Konsistenz als Ü identifiziert. Nur 8% der Antworten entfielen hier auf die Kategorie Ö, ebensoviel wie auf die Kategorie U.

---

<sup>94</sup>Bei einer Präsentation der Versuchsergebnisse im Rahmen einer Postersession der *Interspeech 2013* in Lyon bewerteten jedoch 21 Besucher – in Unkenntnis der Bewertungen in den Perzeptionsexperimenten – diesen Stimulus spontan als Ä-ähnlich. Lediglich eine weitere – frankophone – Besucherin wich hierbei ab, indem sie den Stimulus als nasalisierten, zentralisierten E-Vokal beschrieb.

Die Ergebnisse für die Kategorie IIIb scheinen in dieselbe Richtung zu deuten (Tab. 6.15): Die Stimuli für G und  $c^0$  werden anders als in Kategorie I und den vorangehenden Perzeptionsexperimenten stark divers bewertet. Hier scheint ein Widerspruch zwischen den Formantwerten und einer relativ dazu gesehen zu niedrigen  $F_0$  zu bestehen, der eine sichere Identifikation nicht möglich macht. Bemerkenswert ist die Bewertung der Stimuli für  $c^1$  und  $g^1$ : Diese werden zwar – wie auch in Kategorie I – ebenfalls divers bewertet (mit vergleichbaren Schwerpunkten), doch ist die Streuung um diese Schwerpunkte eine vollständig andere als in Kategorie I. Die Bewertungen für  $c^2$ ,  $g^2$  und  $c^3$  sind im Vergleich zur Kategorie I vollständig gekippt.

Die zu Kontrollzwecken eingesetzten acht Stimuli der Kategorie II wurden erwartungsgemäß mit grundsätzlich sehr hohen Erkennungsraten identifiziert. Im Bereich der Vorderzungenvokale, insbesondere des / $\epsilon$ /, sorgt eine relativ hohe Verwechslungsrate mit dem jeweils nächst geschlosseneren Vokal für geringere Erkennungsraten als im Hinterzungsbereich (Tab. 6.13).

Die hinsichtlich Grundfrequenz und Formanten  $F_1$ – $F_3$  theoretisch identischen Stimuli der Kategorien I und IIIa fallen in ihrer Bewertung vollständig unterschiedlich aus. Mit Ausnahme der höchsten drei Töne wurden alle Stimuli der Kategorie IIIa anderen Vokalqualitäten zugeordnet als ihre Pendants der Kategorie I. Besonders augenfällig ist dies beim Stimulus C, der statt mit  $\ddot{O}$  oder O nun sehr konsistent mit A identifiziert wird. Ähnlich verhält es sich beim Stimulus G, bei dem die Qualität  $\ddot{O}$  zugunsten ihres ungerundeten (Quasi-)Pendants  $\ddot{A}$  aufgegeben wird. Bei den Stimuli  $c^0$ ,  $g^0$  und  $c^1$  weisen jeweils die Vokalqualitäten, die in Kategorie I die höchste Zustimmungsraten erhalten hatten, außerordentlich niedrige Werte auf.

**Tabelle 6.6:** Antworten in Prozent für die Stimuli aus Amorbach im Perzeptionsexperiment Nr. 1. Die jeweils häufigste Antwort ist fett gedruckt, die Grautöne korrespondieren mit den Häufigkeitsklassen (100–80% dunkelgrau, 79–60%, 59–40%, 39–20% (hellstes Grau), 19–0% (keine Schattierung). Alternativ zur Auswahl einer Vokalkategorie war auch die Angabe »kein Vokal« (k. V.) möglich.

PE 1: VH (Amorbach)														
TON	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	k. V.	/i/	/y/	/e/	/ɛ/	/ø,œ/	/a,ɑ/	/ɔ,o/	/u/	Σ
C	66	1036	1628	2856	20	0	0	0	5	55	10	10	0	100
G	98	1283	1952	2973	10	0	0	5	5	70	5	5	0	100
c0	132	585	1545	2260	10	0	0	5	30	30	10	15	0	100
g0	198	742	1752	2641	10	5	0	15	40	20	5	0	5	100
c1	263	754	1758	2583	15	0	0	5	25	30	10	15	0	100
c1	263	859	1738	2639	15	10	0	25	15	20	5	10	0	100
g1	395	825	1992	3040	10	30	0	45	0	10	0	5	0	100
c2	527	1028	2091	2628	40	20	0	25	0	0	10	5	0	100
g2	790	1032	2333	3155	20	45	5	25	0	0	0	0	5	100
c3	1054	1053	2107	3163	15	70	15	0	0	0	0	0	0	100

**Tabelle 6.7:** Antworten in Prozent für die Stimuli aus Meisenheim im Perzeptionsexperiment Nr. 1. Grafische Nuancierung wie bei Tab. 6.6.

PE 1: VH (Meisenheim)														
TON	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	k. V.	/i/	/y/	/e/	/ɛ/	/ø,œ/	/a,ɑ/	/ɔ,o/	/u/	Σ
C	70	1012	1548	2550	30	0	0	5	5	30	25	5	0	100
G	105	726	1395	2122	10	0	0	0	20	60	10	0	0	100
c0	141	816	1519	2289	30	0	0	15	20	30	5	0	0	100
g0	210	803	1824	2654	15	0	0	10	45	10	5	15	0	100
c1	281	1063	1994	2997	35	0	5	20	15	10	10	0	5	100
g1	412	1274	2052	3293	30	30	0	0	10	5	15	0	10	100
c2	562	1299	2149	2785	30	15	0	35	0	0	20	0	0	100
g2	844	845	1719	2532	10	60	10	15	0	0	5	0	0	100
c3	1124	1123	2247	2495	20	70	5	5	0	0	0	0	0	100



**Tabelle 6.8:** Antworten in Prozent für die Stimuli aus Simmern im Perzeptionsexperiment Nr. 1. Grafische Nuancierung wie bei Tab. 6.6.

PE 1: VH (Simmern)														
TON	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	k. V.	/i/	/y/	/e/	/ɛ/	/ø,œ/	/a,ɑ/	/ɔ,o/	/u/	Σ
C	69	982	1682	2823	40	0	0	0	0	25	10	25	0	100
G	102	644	1304	1957	5	0	0	0	10	80	0	0	5	100
c0	136	773	1426	2174	5	0	0	5	35	35	10	10	0	100
g0	205	754	1661	2580	20	0	0	20	35	10	10	5	0	100
c1	274	871	1899	1919	15	0	0	15	25	15	20	5	5	100
g1	408	852	2029	2500	20	0	0	35	20	0	20	5	0	100
c2	548	1104	2011	2430	25	5	5	20	5	0	35	0	5	100
g2	818	911	2234	2557	30	50	0	10	0	0	5	0	5	100
c3	1093	1092	2185	2912	15	80	0	0	0	0	5	0	0	100

**Tabelle 6.9:** Antworten in Prozent für die Stimuli aus Waltershausen im Perzeptionsexperiment Nr. 1. Grafische Nuancierung wie bei Tab. 6.6.

PE 1: VH (Waltershausen)														
TON	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	k. V.	/i/	/y/	/e/	/ɛ/	/ø,œ/	/a,ɑ/	/ɔ,o/	/u/	Σ
C	69	448	1585	2573	15	0	0	0	5	40	15	20	5	100
G	103	474	998	2040	10	0	0	0	0	20	30	35	5	100
c0	139	594	1673	2739	20	0	0	25	15	15	5	10	10	100
g0	208	437	1573	2114	15	0	5	10	0	65	0	0	5	100
c1	277	808	1728	2959	10	0	5	10	5	10	35	5	20	100
g1	415	1218	2071	2878	25	10	15	20	5	5	20	0	0	100
c2	554	575	1690	2301	35	20	35	5	0	0	5	0	0	100
g2	832	831	1663	2499	35	15	35	0	0	0	5	0	10	100
c3	1109	1110	1903	2218	30	35	5	0	5	5	15	0	5	100

**Tabelle 6.10:** Antworten in Prozent für die Stimuli aus Simmern im Perzeptionsexperiment Nr. 2. Grafische Nuancierung wie bei Tab. 6.6.

PE 2: VH, CR und TR (Simmern)														
REG.	TON	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	i	ü	e	ä	ö	a	o	u	Σ
VH	C	69	982	1682	2823	0	2	2	8	31	7	38	11	100
	G	102	644	1304	1957	0	1	2	8	85	2	1	0	100
	c0	136	773	1426	2174	0	1	14	15	64	3	2	0	100
	g0	205	754	1661	2580	0	0	15	36	41	7	1	0	100
	c1	274	871	1899	1919	2	2	20	21	29	20	3	3	100
	g1	408	852	2029	2500	5	6	59	16	6	6	0	3	100
	c2	548	1104	2011	2430	15	13	37	2	8	23	1	1	100
	g2	818	911	2234	2557	51	7	25	2	3	7	1	3	100
	c3	1093	1092	2185	2912	84	9	3	0	0	3	0	0	100
CR	C	69	1097	1688	3014	0	0	3	13	28	30	17	9	100
	g0	205	941	1434	1989	1	0	28	20	34	14	3	0	100
TR	C	69	695	1494	2001	0	1	1	6	20	14	33	25	100
	g0	205	1576	1747	2631	6	16	22	2	13	10	8	23	100

**Tabelle 6.11:** Antworten in Prozent für die Stimuli aus Waltershausen im Perzeptionsexperiment Nr. 2. Grafische Nuancierung wie bei Tab. 6.6.

PE 2: VH (Waltershausen)													
TON	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	i	ü	e	ä	ö	a	o	u	Σ
C	69	453	1587	2573	0	1	1	2	33	1	40	21	100
G	103	529	996	2040	1	1	6	3	40	9	32	7	100
g0	208	434	1567	2114	1	26	15	1	32	3	9	11	100
c1	277	843	1838	2959	2	9	10	2	10	5	5	56	100
g1	415	1286	2061	2879	9	29	28	2	13	6	5	9	100
c2	554	578	1690	2301	33	21	28	1	7	2	0	8	100
g2	832	831	1663	2499	39	22	3	1	1	8	3	22	100
c3	1109	1110	1964	2218	75	11	2	1	0	9	0	1	100

**Tabelle 6.12:** Antworten in Prozent für die originalen VH-Stimuli aus Simmern im Perzeptionsexperiment Nr. 3 (Kategorie I). Grafische Nuancierung wie bei Tab. 6.6.

PE 3-I: VH (Simmern) – Originaltöne													
TON	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	i	ü	e	ä	ö	a	o	u	Σ
C	69	982	1682	2823	0	3	6	4	48	5	24	10	100
G	102	644	1304	1957	0	3	5	4	83	2	1	2	100
c0	136	773	1426	2174	2	5	16	11	65	1	0	0	100
g0	205	754	1661	2580	2	10	21	23	41	1	0	2	100
c1	274	871	1899	1919	4	7	18	38	24	3	2	4	100
g1	408	852	2029	2500	28	7	47	11	3	0	2	2	100
c2	548	1104	2011	2430	36	2	23	10	2	21	3	3	100
g2	818	911	2234	2557	67	9	10	6	1	6	0	1	100
c3	1093	1092	2185	2912	86	5	5	0	1	2	0	1	100

**Tabelle 6.13:** Antworten in Prozent für die synthetisierten Langvokale nach [Simpson 1998] im Perzeptionsexperiment Nr. 3 (Kategorie II). Grafische Nuancierung wie bei Tab. 6.6.

PE 3-II: Synthetisierte Vokale (Klatt)													
VOK.	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	i	ü	e	ä	ö	a	o	u	Σ
/u/	100	347	1060	2454	0	0	0	0	2	0	1	97	100
/o/	100	420	955	2538	0	0	0	0	0	1	87	12	100
/ɔ/	100	571	1155	2491	0	0	0	0	0	6	94	0	100
/a/	100	716	1296	2589	0	2	1	11	7	78	1	0	100
/ø/	100	409	1538	2308	0	8	0	4	86	0	2	0	100
/ɛ/	100	515	1972	2467	0	0	42	55	3	0	0	0	100
/e/	100	381	2068	2601	10	5	83	0	3	0	0	0	100
/i/	100	311	2156	2711	69	8	24	0	0	0	0	0	100

**Tabelle 6.14:** Antworten in Prozent für die resynthetisierten VH-Stimuli aus Simmern im Perceptionsexperiment Nr. 3 (Kategorie IIIa). Grafische Nuancierung wie bei Tab. 6.6.

PE 3-IIIa: VH (Simmern) – Resynthese (Klatt)													
TON	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	i	ü	e	ä	ö	a	o	u	Σ
C	69	982	1682	2823	0	1	0	14	4	80	1	0	100
G	102	644	1304	1957	0	4	1	41	20	33	1	0	100
c0	136	773	1426	2174	0	3	2	34	5	56	0	0	100
g0	205	754	1661	2580	0	5	11	67	10	3	1	3	100
c1	274	871	1899	1919	15	18	30	15	15	1	2	4	100
g1	408	852	2029	2500	12	7	30	28	10	9	2	3	100
c2	548	1104	2011	2430	29	10	19	16	10	13	1	2	100
g2	818	911	2234	2557	59	15	9	5	4	4	1	4	100

**Tabelle 6.15:** Antworten in Prozent für die synthetisierten Stimuli mit konstanter F<sub>0</sub> und F<sub>1</sub>-F<sub>3</sub> nach VH (Simmern) im Perceptionsexperiment Nr. 3 (Kategorie IIIb). Grafische Nuancierung wie bei Tab. 6.6.

PE 3-IIIb: F <sub>0</sub> =100, F <sub>1</sub> -F <sub>3</sub> nach VH (Simmern)													
TON	F <sub>0</sub>	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	i	ü	e	ä	ö	a	o	u	Σ
C	100	982	1682	2823	0	1	6	12	5	67	9	1	100
G	100	644	1304	1957	0	4	7	24	19	16	30	0	100
c0	100	773	1426	2174	0	3	4	25	10	32	27	0	100
g0	100	754	1661	2580	0	3	6	51	10	12	15	2	100
c1	100	871	1899	1919	0	4	10	32	10	20	22	2	100
g1	100	852	2029	2500	0	4	7	45	3	22	18	2	100
c2	100	1104	2011	2430	0	4	9	12	16	50	6	3	100
g2	100	911	2234	2557	0	2	10	36	6	27	19	1	100
c3	100	1092	2185	2912	0	3	7	23	19	41	8	0	100

**Tabelle 6.16:** Antworten in Prozent für die synthetisierten Stimuli mit  $F_0$  nach VH (Simmern) und  $F_1$ - $F_3$  nach [Simpson 1998] im Perzeptionsexperiment Nr. 3 (Kategorie IIIc). Es wurden pro Stimulus die Formantwerte des Vokals verwendet, der im Perzeptionsexperiment Nr. 2 mit der jeweiligen Tonhöhe ( $F_0$ ) am häufigsten assoziiert wurde. Grafische Nuancierung wie bei Tab. 6.6.

PE 3-IIIc: $F_0$ VH (Simmern), $F_1$ - $F_3$ nach Simpson (1998)													
TON VOK.	$F_0$	$F_1$	$F_2$	$F_3$	i	ü	e	ä	ö	a	o	u	$\Sigma$
g2	/i/	818	311	2156	2711	87	6	2	0	4	1	0	100
c2	/e/	584	381	2068	2601	65	20	8	1	3	1	0	100
gl	/e/	408	381	2068	2601	70	23	4	0	1	1	0	100
cl	/ø/	274	409	1538	2308	5	79	1	0	8	0	0	100
g0	/ø/	205	409	1538	2308	3	42	0	0	52	0	1	100
c0	/ø/	136	409	1538	2308	3	24	1	0	71	1	0	100
G	/ø/	102	409	1538	2308	3	14	0	0	83	0	0	100
C	/o/	69	420	955	2538	0	0	0	0	1	0	89	100

## 6.4 Diskussion

Die kritische Gegenüberstellung von historischen »Ohrenzeugenberichten« über die Klangqualität der VH und aktuellen akustischen und perzeptorischen Versuchen konnte zweierlei zeigen: Zum einen, dass zwar – wie angesichts der teilweise geradezu euphorischen zeitgenössischen Berichte anders hätte vermutet werden können – offenbar nur sehr bedingt eine Beziehung zwischen dem Klang einer typischen VH des 18. Jahrhunderts und menschlichen Vokalen hergestellt werden kann. Zum anderen wurde aber auch deutlich, dass die angeführten früheren Analysen zu diesem Thema keine verlässliche Aussagekraft besitzen, was nicht zuletzt der seinerzeit noch erheblich beschränkteren spektralen Analysemöglichkeiten geschuldet sein dürfte.

Bemerkenswert ist hierbei die Parallele zwischen den Sprachsynthesen KEMPELENS bzw. den Synthesen in dessen Tradition und der frühen Benennung der *Regalregister*. Diese wurden insbesondere im Bereich der franko-flämischen Orgelbautradition, aber auch in Nordwestdeutschland im 16. und 17. Jahrhundert häufig parallel als *Vox humana* und als als *Jeu d'enfant* oder gar *Kinderbass* bezeichnet [Eberlein 2009, 701]. Nimmt man diese Benennungen als Charakterisierungen ernst, so sollte man annehmen können, dass gerade die VH-Töne derjenigen Tonlage dieser *Register*, die im Stimmambitus eines singenden Knaben liegen, als besonders stimm-

bzw. vokalähnlich bewertet werden. Die hier behandelten Perzeptionsexperimente bestätigen dies jedoch nicht sondern deuten im Gegenteil darauf hin, dass gerade diese mittlere Tonlage als wenig eindeutig wahrgenommen wird.

Ungeachtet dieser fehlenden Vokalähnlichkeit, die auch den Hörern des 18. Jahrhunderts evident gewesen sein muss, scheint aber beispielsweise bei den Gebrüdern STUMM grundsätzlich ein gewisser Anspruch bestanden zu haben, ihre *VH-Register* so zu gestalten, dass ihr Klang halbwegs menschnah geriet. Dafür spricht jedenfalls die Registerbezeichnung »Menschenstimme«, die von diesen Orgelbauern auf die tiefste Pfeife C in Amorbach und Simmern geritzt wurde [vgl. Eppelsheim 1971, 47, Fußnote 18 und Bösen 1981, 90].<sup>95</sup> Der Registername wurde hier also offenbar nicht als *Terminus technicus* verstanden, sondern als eine programmatische Klangbeschreibung, die man ohne Bedenken ins Deutsche übertrug.<sup>96</sup>

Die hier vorgestellten Ergebnisse der akustischen Analysen und der Perzeptionstests legen jedoch nahe, dass zwischen dem Klang einer *VH* und dem von menschlichen Vokalen eine zumindest graduell größere Ähnlichkeit besteht als es bei anderen *Lingualregistern* der Fall ist. Hierfür dürfte in der Tat der bereits von LOTTERMOSER angeführte Umstand verantwortlich sein, dass die Abmessungen der *VH-Resonatoren*, anders als bei eigentlich allen anderen *Lingualregistern*, nicht in direktem proportionalem Verhältnis zur Tonhöhe der jeweiligen Pfeife stehen. Dass die von LOTTERMOSER festgestellten Formantstrukturen grundsätzlich jedoch kein Alleinstellungsmerkmal der *VH* und ihrer charakteristischen Resonatorgestaltung sind, konnte ebenso gezeigt werden. Die Strukturen der untersuchten *VH* zeichnen sich allerdings gegenüber den kontrastiv untersuchten *Registern Trompette* und *Cromorne* durch klar voneinander abgegrenzte Frequenzbänder aus, wie sie ähnlich auch für menschliche Vokale charakteristisch sind. Ziel künftiger Untersuchungen zu diesem Aspekt muss es unter anderem sein, eine größere Zahl von historischer *Zungenregistern* und zudem auch charakteristische *Labialregister* auf mögliche energiereiche Frequenzbänder, die als Formanten angesprochen werden könnten, zu untersuchen.

Beim Vergleich der Formanten der vier *VH* mit denen von Referenzsprechern ist die beinahe lineare Anordnung der einzelnen Töne eines *Registers* zueinander auffällig (vgl. Abb. 6.11 und 6.12 auf S. 141)). Bei flüchtiger Betrachtung der Formantwerte scheinen beispielsweise die Werte für  $F_2$

<sup>95</sup>Ein fotografischer Beleg wäre aufgrund der sehr ungünstigen Kontrastverhältnisse nur mit sehr aufwändiger Streiflicht-Technik möglich gewesen und musste daher leider unterbleiben.

<sup>96</sup>Die Bezeichnung am Spieltisch lautet in beiden Fällen dann regulär »Vox humana«.

**Tabelle 6.17:** Werte für  $F_1$ – $F_3$  sowie doppelte und dreifache  $F_1$ -Werte für die VH in Simmern und Waltershausen.

Ton	Simmern					Waltershausen				
	$F_1$	$F_1*2$	$F_2$	$F_1*3$	$F_3$	$F_1$	$F_1*2$	$F_2$	$F_1*3$	$F_3$
C	982	1964	1628	2946	2823	453	906	1587	1359	2573
G	664	1382	1304	1992	1957	529	1058	996	1587	2040
$c^0$	773	1546	1426	2319	2174	576	1152	1670	1728	2739
$g^0$	754	1508	1661	2262	2580	434	868	1567	1302	2114
$c^1$	871	1742	1899	2613	1919	843	1686	1838	2529	2959
$g^1$	852	1704	2029	2556	2500	1286	1572	2061	3858	2878
$c^2$	1104	2208	2011	3312	2430	578	1156	1690	1734	2301
$g^2$	911	1822	2234	2733	2557	831	1662	1663	2493	2499
$c^3$	1092	2184	2185	3276	2912	1110	2220	1964	3330	2218

in etwa das Doppelte des jeweiligen  $F_1$ -Werts zu betragen. Eine solche unidirektionale Formantprogression im Vokalraum der VH wäre erklärbar durch ihre Konstruktion: Als einziger Proportionsparameter verändert sich bei den VH-Pfeifen unterschiedlicher Tonhöhe das Längenmaß der *Resonatoren* (wenn auch nicht linear-proportional zur Tonhöhe), während alle anderen Maße weitgehend konstant bleiben. Hinzu kommt, dass die Resonatoren stets nur eine Öffnung besitzen (an ihrer Oberseite) und in ihrem Verlauf keinerlei Verengungen oder anderweitige Querschnittsveränderungen aufweisen. Demgegenüber wird für menschliche Vokale gemeinhin angenommen, dass sie einen vorderen und einen hinteren Resonanzraum besitzen. Für einen Resonanzraum wie den für die VH beschriebenen sollte man annehmen, dass die höheren Formanten jeweils Vielfache des  $F_1$  darstellen. Tab. 6.17 zeigt jedoch exemplarisch für die *Register* in Simmern und Waltershausen, dass dies für die untersuchten VH nicht der Regelfall ist. Einzig im Falle des Tons  $g^2$  für Waltershausen und mit gewissen Einschränkungen für die Töne C, G und  $c^3$  aus Simmern trifft diese Hypothese zu.

Die Formantmessungen der untersuchten VH zeigen einen kleineren Vokalraum mit deutlich erhöhten Formantwerten gegenüber einer menschlichen Referenz-Sprechstimme (vgl. Abb. 6.11 und 6.12 auf S. 141). Der Grund hierfür dürfte in dem vergleichsweise kleinen »Vokaltrakt« der Orgelpfeifen zu suchen sein. Dieser beträgt beispielsweise für den Ton  $g^0$  bei allen drei STUMM-*Registern* 14,0 x 2,7 cm. POMPINO-MARSCHALL gibt die durchschnittlichen Maße eines Vokaltrakts eines männlichen Referenzsprechers mit 17,0 x 4,5 cm an [Pompino-Marschall 2009, 160]. Abweichend

hiervon nennen FITCH & GIEDD Längenmaße von lediglich 15,0 cm für junge erwachsene männliche Sprecher (19–25 Jahre) und 14,0 cm für jugendliche männliche Sprecher (13–16 Jahre), was auffällig mit den für die VH mitgeteilten Maßen korreliert [Fitch & Giedd 1999]. Dies ist insbesondere vor dem Hintergrund bemerkenswert, dass es in der Regel mit jugendlichen männlichen Sängern besetzte Chöre waren, deren Klang in der Ursprungszeit der *Vox humana*-Registrierung durch diese substituiert werden sollte.

Alle drei Perzeptionstests konnten zeigen, dass zwar nicht alle, sehr wohl jedoch viele der dargebotenen VH-Stimuli von der Mehrheit der Versuchspersonen konsistent als vokalähnlich eingestuft wurden. Dies gilt insbesondere für die hohen und tiefen Töne und hier in besonderer Konsistenz für die Stimuli aus Simmern. In den mittleren Lagen war dagegen eine deutlich stärkere Streuung zu beobachten. Die Erkennungsraten für die peripheren Töne in Simmern weisen eine deutliche Ähnlichkeit zu Verwechslungstests mit CV- bzw. VC-Silben im Englischen auf, wie sie von WEBER & SMITS berichtet werden [Weber & Smits 2003]. Dort lagen die Erkennungsraten der Vokale teilweise sogar bei unter 45%.

Die von Versuchsperson zu Versuchsperson stark schwankenden Dauern im ersten Experiment lassen den Schluss zu, dass die Aufgabenstellung bisweilen eine nicht unerhebliche Herausforderung darstellte. Eine zusätzliche Abfrage zur empfundenen Synthese-Qualität der jeweiligen Stimuli in allen drei Experimenten hätte einen Hinweis darauf geben können, als wie stark vokalähnlich der jeweilige Stimulus überhaupt empfunden wurde.

Insbesondere die Ergebnisse des dritten Perzeptionsexperiments lassen den Schluss zu, dass zumindest im Bereich synthetischer Vokalklänge die Grundfrequenz eine erhebliche Rolle für die Perzeption der Vokalqualität spielt. Der Bewertungskontrast zwischen den synthetisch erzeugten Stimuli der Kategorien IIIa–c und denen der Kategorie I (bzw. der vorangehenden beiden Perzeptionstests) zeigt zudem, dass Grundfrequenz und Formanten nicht unabhängig voneinander sind, sondern – wiederum zumindest für diesen Bereich – aufeinander aufbauen. Allerdings zeigt insbesondere der Kontrast zwischen den Kategorien I und IIIa auch, dass ganz offenbar nicht nur die Faktoren Grundfrequenz und Formantwerte eine Rolle in der Perzeption von Vokalqualitäten spielen. Die gegenüber ihren Pendants der Kategorie I vollständig andere – und oft inkonsistentere – Bewertung der Stimuli der Kategorie IIIa legt den Schluss nahe, dass einerseits »unharmonische« Faktoren für die Kategorisierung der originalen VH-Stimuli



eine Rolle spielen müssen und dass andererseits die verwendete Synthese nicht in allen Aspekten zufriedenstellend arbeitete.<sup>97</sup>

Ein Vergleich der Testergebnisse mit den Befunden von EULER und WILLIS zeigt deutliche Parallelen auf: EULER hatte den Klang der VH als Ä- bzw. Ö-ähnlich beschrieben [Euler 1768, 97] (vgl. Zitat auf S. 24), WILLIS hatte die »Vokalqualität« eines VH-Tons unmittelbar mit dessen Tonhöhe in Verbindung gebracht: Bis zu einer maximalen Tonhöhe von 65 Hz könne eine VH O-ähnlich klingen, bis 175 Hz wie ein A und bis 290 Hz wie ein E [Willis 1832, 432] (vgl. Zitat auf S. 24). Diese – wohl kaum empirisch gewonnenen – Befunde lassen sich zumindest in ihrer grundsätzlichen Tendenz gut mit den hier präsentierten Ergebnissen vergleichen.

Auffällig ist die unterschiedliche Bewertung ein und desselben Tones für die drei doch vermeintlich baugleichen VH-Register der Gebrüder STUMM unter kontrolliert vergleichbaren Aufnahmebedingungen. So wurde beispielsweise der Ton G in Simmern in allen drei Perzeptionsexperimenten zwischen 80% und 85% als Ö bewertet, in Amorbach und Meisenheim im ersten Perzeptionsexperiment hingegen nur mit 70% bzw. 60%. Eine Erklärung hierfür könnte die deutlich unterschiedliche Intensitätsverteilung der Obertöne in allen drei Registern sein. Diese wiederum kann mannigfaltige Ursachen haben. Die zwei wahrscheinlichsten sind eine geringfügig unterschiedliche Zusammensetzung der Messinglegierung, aus der die Zungenblätter hergestellt wurden und der Einfluss diverser, von Orgel zu Orgel aber unterschiedlicher Nachintonationen im Laufe der Jahrhunderte. Wesentlich ist in diesem Zusammenhang der Einfluss der Biegung des Zungenblattes. Diese ist für den Klang einer Lingualpfeife in hohem Maße mitverantwortlich. Dass ein Zungenblatt über eine derart lange Zeit seine ursprüngliche Biegung behalten hat, darf ebenso ausgeschlossen werden wie eine identische Biegung bei allen drei Registern (vgl. [Goebel 1975]).

Obwohl die Stimuli aus Waltershausen durch die Kombination mit dem Register *Gedeckt 8'* theoretisch der realistischen Klangpräsentation einer VH näher kamen, wurden diese Stimuli durchgehend inkonsistenter bewertet als ihre Pendanten besonders aus Simmern. Möglicherweise spielte hier die gegenüber den übrigen VH durch das *Labialregister* deutlich verstärkte Grundfrequenz eine Rolle. Die höheren Teiltöne besitzen in Waltershausen eine erheblich niedrigere Intensität als bei den anderen drei VH.

Die Grundfrequenz allein vermag aber die Zuordnung der einzelnen Stimuli zu Vokalqualitäten nicht befriedigend zu erklären. Hierfür bedarf es der Beiziehung der Formantwerte. So zeigt der Ton c<sup>3</sup> (mit I assoziiert) aus Simmern beispielsweise neben seiner hohen F<sub>0</sub> auch sehr hohe Werte

<sup>97</sup>Es bliebe zu untersuchen, in wie weit die Tatsache der höheren F<sub>0</sub> als F<sub>1</sub> bei einigen Stimuli der Kategorie IIIc zu deren perzeptiver Wahrnehmung beigetragen hat.

für  $F_2$  und  $F_3$ , wohingegen der Ton G (schwerpunktmäßig assoziiert mit Ö) die absolut tiefsten Werte für  $F_2$  und  $F_3$  zeigt.

Die gegenüber den eigentlichen Perzeptionsexperimenten bemerkenswert konsistente qualitative Bewertung des VH-Stimulus  $c^1$  aus Simmern durch Besucher im Rahmen der *Interspeech 2013* ist möglicherweise damit zu erklären, dass es sich einerseits weitestgehend um Fachpublikum mit Vorbildung im Bereich der Sprachsynthese handelte und andererseits allen der Stimulus mit demselben Paar Kopfhörer präsentiert wurde. Gegen letzteren Grund spricht indessen, dass sich dann auch bei den übrigen Stimuli eine ähnliche Streuung ergeben haben müsste.

Die zur Kontrolle eingesetzten Aufnahmen der historischen *Lingualregister Trompette* und *Cromorne* zeigen ebenfalls deutliche Formantstrukturen. Die entsprechenden Stimuli wurden allerdings weit weniger konsistent einer bestimmten Vokalqualität zugeordnet als es bei den jeweiligen VH-Äquivalenten der Fall war.

Offenbar haben unabhängig von der vermutlich vollkommen anderen Gesangstradition in der Blütezeit der VH, mit deren Klangwirkung dieses Orgelregister verglichen werden konnte, auch andere Faktoren zur Attributierung des VH-Klanges als sprachähnlich beigetragen. Der Klang der VH konnte durchaus als vokalähnlich wahrgenommen werden.

Der amerikanische Arzt und Schriftsteller OLIVER WENDELL HOLMES (1809-1894) stellt pointiert fest:

»As to the vox humana, which is so celebrated, it does not all resemble a human voice, though a very good stop of the kind; but the world is very apt to be imposed upon by names; the instant a common hearer is told that an organist is playing upon a stop which resembles the human voice, he supposes it to be very fine and never inquires into the propriety of the name, or exactness of the imitation. However, with respect to our own feelings, we must confess, that, of all stops which we have yet heard, that have been honored with the appellation of vox humana, no one in the treble part has ever reminded us of anything, so much as the cracked voice of an old woman of ninety, or, in the lower parts, of Punch singing through a comb.« [Holmes 1864, 394 f.]

Einen direkten Bezug zwischen der VH und einer konkreten sprachsynthetischen Anwendung stellte die Preisfrage der St. Petersburger Akademie der Wissenschaften für das Jahr 1780 dar (vgl. Kapitel 3.2).

Die in diesem Zusammenhang entstandene »Vokalorgel« Chritian Gottlieb Kratzensteins stellt einen bemerkenswerten Spagat zwischen musikalischer und sprachsynthetischer Anwendungsmöglichkeit dar. Im folgenden Kapitel soll es um die Frage gehen, in wie weit dieses Vokalsynthesekonzept heute rekonstruierbar ist.



# Kapitel 7

## Kratzensteins »Vokalorgel«

### 7.1 Die durchschlagende Lingualpfeife

Auf den Kontext von KRATZENSTEINS Arbeit zur Vokalsynthese wurde bereits im Kapitel 3.2 wurde ausführlich eingegangen. Zum tieferen Verständnis seines Anregungskonzeptes seien hier diejenigen Passagen seines »*Tentamen*« zusammenfassend referiert, die die Konstruktion der »Sprachorgel« illustrieren (vgl. hierzu auch die Abb. 7.1 und 7.2). Die wörtlichen Übersetzungen der entsprechenden Textabschnitte finden sich im Appendix ab S. LVII.

In § 15 wird zunächst eine Vorrichtung erläutert, die die von KRATZENSTEIN postulierte Funktionsweise des Kehldeckels darstellt (vgl. Abb. 7.1 oben links). Dieser wird hier durch die bewegliche Membrane am oberen Ende einer Röhre simuliert. Die Membrane soll elastisch befestigt sein, so dass sie durch die von unten einströmende Luft horizontal in Schwingung versetzt werden kann. Wesentlich ist hierbei, dass das als Membran fungierende Plättchen exakt in den Innendurchmesser der Röhrenmündung passen muss. Mit dieser Konstruktion ist das Funktionsprinzip einer *durchschlagenden Zungenpfeife* vermutlich erstmals in Europa angewendet worden.

Das hier beschriebene Instrument entspricht jedoch, wie auch die bei KRATZENSTEIN beigegebene Abbildung zeigt, in keiner Weise einer *Zungenpfeife*, wie sie aus dem Orgelbau bekannt ist. Im weiteren Verlauf schildert KRATZENSTEIN in § 24 die Möglichkeit, herkömmliche *Zungenpfeifen* in ihrer Konstruktion so zu modifizieren, dass sie der zuvor geschilderten Konstruktion in ihrem Klangergebnis zumindest nahekommen. Kernstück hierbei ist die sog. *Fütterung* der *Kehle*, die im Orgelbau von alters her

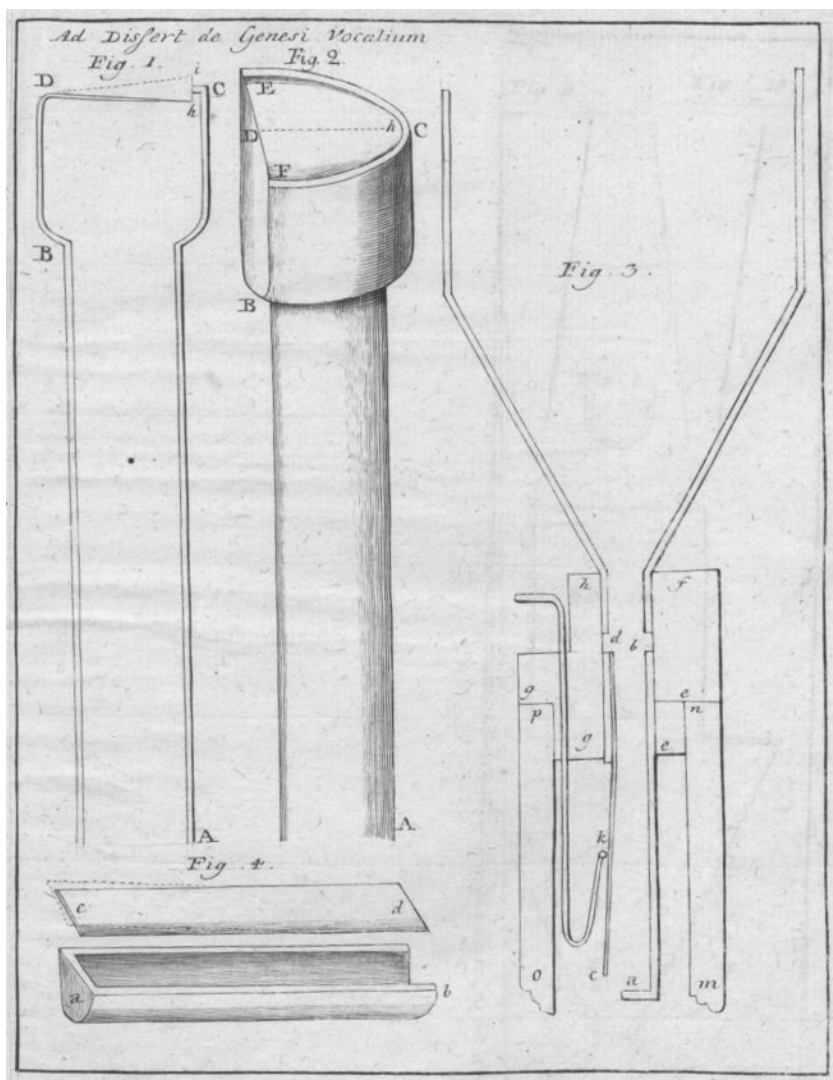
ebenso bekannt ist wie auch das erwähnte Versehen der *Kehlenoberseite* mit einer Auflage aus Zinn. Bemerkenswert indes sind die Anregungen zur Verwendung von organischen Materialien für *Zungenblätter*. Im folgenden § 25 erläutert KRATZENSTEIN sodann sein Konzept der *durchschlagenden Zungenpfeifen*, das mehr noch als die zuvor genannten Verbesserungsvorschläge geeignet sei, Klänge ähnlich der menschlichen Stimme hervorzu- bringen. Erstaunlich ist hier die Aussage, dass die Klangqualität einer solchen Pfeife – ohne Resonator benutzt – einem Diphthong ae entspreche (»sine tuba adjuncta edit diphtongum ae« [Kratzenstein 1781, 38]) und erst durch Einsatz verschieden geformter Resonatoren verschiedene monophthongale Vokalqualitäten hervorgebracht werden könnten.<sup>98</sup>

## 7.2 Zur möglichen Rekonstruktion der »Vokalorgel«

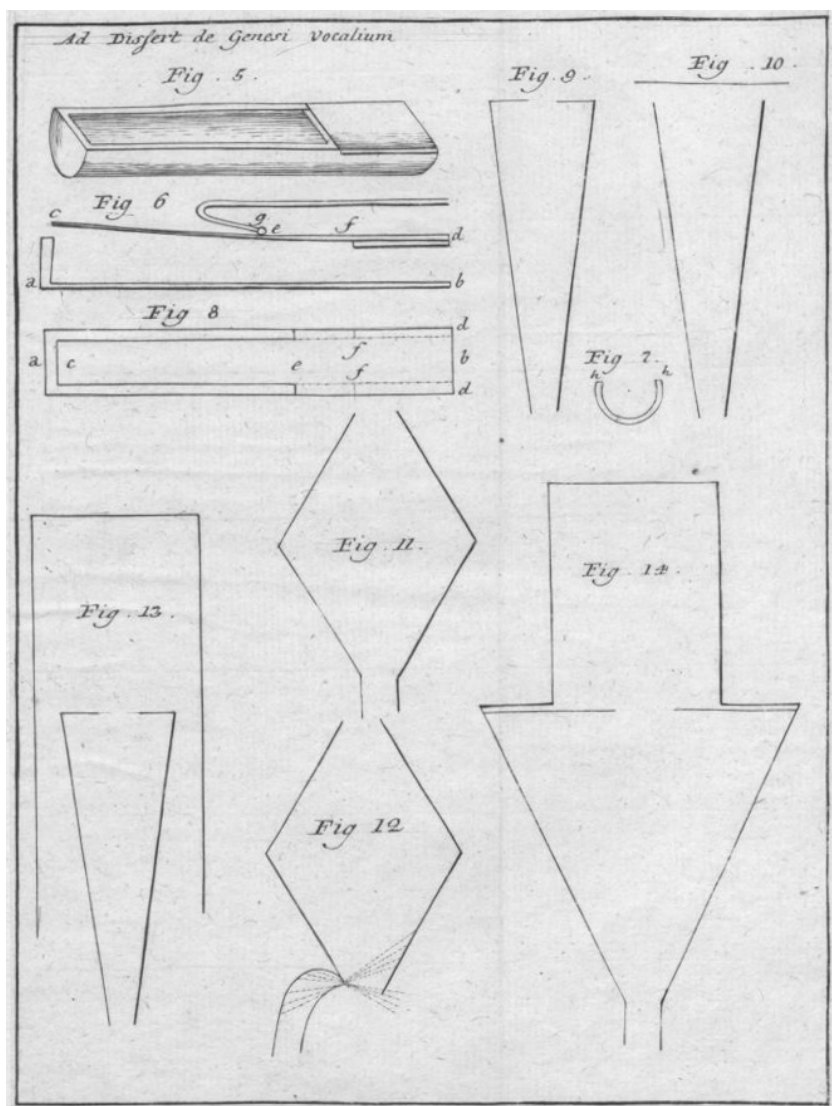
Ist eine Rekonstruktion von KEMPELENS Sprachmaschine dank seiner recht detaillierten Beschreibung vergleichsweise einfach, so ergibt sich bei einem ähnlichen Projekt bezüglich der KRATZENSTEINISCHEN Synthese eine Vielzahl grundlegender Probleme. Wohl hatte KRATZENSTEIN die grundsätzliche Konstruktion seiner Pfeifen beschrieben und sogar schematische Querschnitte der *Resonatoren* mitgeteilt, doch als Grundlage einer Rekonstruktion kann all dies nur sehr eingeschränkt verwendet werden. Insbesondere das vollständige Fehlen von Maßangaben oder zumindest konkreten Proportionen erschwert die Konstruktion erheblich.

Fraglich sind beispielsweise auch die Proportionen der Resonatoren untereinander. Während [YOUNG 1845, PL. XXVI] alle fünf als gleich groß darstellt (vgl. Abb. 3.3 auf S. 44), verhält es sich bei KRATZENSTEIN selbst anders. Hier sind die Exemplare für O und U erheblich größer als die übrigen dargestellt (vgl. Abb. 3.3 auf S. 44 und 7.2). Dies ist zwar vermutlich der ökonomischen Platzausnutzung auf dem Tableau geschuldet, doch kann dies nicht als gegeben vorausgesetzt werden. Weiterhin wissen wir nichts darüber, ob es sich bei KRATZENSTEINS Darstellung eher um skizzenhafte Andeutungen oder um exakt proportionale und maßstabsgerechte Darstellungen handelt. Anhaltspunkte konnte hier JOHN OHALA geben, der vor wenigen Jahren eine Rekonstruktion zweier *Resonatoren* aus starker Pappe versucht hat und zu dem ernüchternden Ergebnis kam, dass nur der Resonator für A eine halbwegs befriedigende Synthese erlaube [Ohala 2011]. Die praktische Umsetzung von KRATZENSTEINS recht vagen Erläuterun-

<sup>98</sup>Vgl. hierzu das Zitat LEONHARD EULERS auf S. 24!



**Abbildung 7.1:** *Tabula I* aus [Kratzenstein 1781]. Fig. 1 und 2 zeigen das in §15 erläuterte Modell eines menschlichen Stimmapparates. Fig. 3 zeigt den Schnitt einer im Orgelbau gebräuchlichen aufschlagenden Zungenpfeife, Fig. 4 die hierzu notwendigen Bauteile Kehle und Zungenblatt.



**Abbildung 7.2:** *Tabula II* aus [Kratzenstein 1781]. Fig. 5 – 8 zeigen verschiedene Ansichten einer nach KRATZENSTEINS Vorgaben modifizierten Kehle einer Zungenpfeife. Fig. 9 – 14 zeigen Schnittzeichnungen der Resonatoren, mit deren Hilfe die Vokale a, e, i, o, u synthetisierbar sein sollen.



gen zur Konstruktion der einzelnen Resonatoren wird nicht ohne einen erheblichen Aufwand zu bewerkstelligen sein. Weder gibt er konkrete Tönehöhen für die verwendeten *Zungenpfeifen* an, noch Maße/Proportionen für *Kehlen*, *Zungenblätter* und *Resonatoren*. Für letztere werden einzig – und auch dies nicht durchgehend und eindeutig – die Proportionen ihrer Bestandteile zueinander beschrieben [Kratzenstein 1782, 376 ff.]. Zudem stellt KRATZENSTEIN beispielsweise für die Synthese des Vokals O gleich vier unterschiedliche Konstruktionsvarianten des *Resonators* vor, von denen sich jedoch nur zwei im Abbildungsteil finden [Kratzenstein 1782, 378 f.].<sup>99</sup> Dies sind ausgerechnet die konstruktiv weniger komplexen Modelle. Bemerkenswert sind KRATZENSTEINS Einlassungen zum Vokal I, bei dem er angibt, die entsprechende Synthesekonstruktion (die anders als die übrigen vier Vokalpfeifen gerade *nicht* mit einer *lingualen* Anregung arbeitet) bereits vor über zehn Jahren entwickelt, ihre Konstruktion aber wieder vergessen zu haben [Kratzenstein 1782, 377]. Das von ihm im »*Tentamen*« präsentierte Konzept entspricht weitestgehend der Tonerzeugung einer Blockflöte. Vor dem Hintergrund der im Kapitel 6.3 vorgestellten Ergebnisse aus Perzeptionstests darf vermutet werden, dass KRATZENSTEIN mit dieser Konstruktion schlicht einen recht hohen flötenartigen Ton erzeugte, der aufgrund eben seiner Höhe als I charakterisiert werden konnte.

In bemerkenswertem Kontrast zur geschilderten Ungenauigkeit der Angaben zur Konstruktion der *Resonatoren* steht KRATZENSTEINS beinahe detailversessene Beschreibung der Konstruktion einer *durchschlagenden Zungenpfeife* nach seiner Erfindung [Kratzenstein 1782, 375 f.]. Viele der von ihm geschilderten Arbeitsschritte sind heute zwar nicht mehr üblich oder gar praktikabel, aber es zeigt sich eine deutlich präzisere, mehr dem seinerzeit hochgelobten Mechanikus KRATZENSTEIN angemessene Herangehensweise.

Nichts ist im Übrigen über den technischen Aufbau und die Steuerung der »Vokalorgel« bekannt. Man könnte vermuten, dass es sich um eine Konstruktion ähnlich der gehandelt haben könnte, die KEMPELEN als seinen zweiten Prototyp abbildet [Kempelen 1791a, Tab. XVII, Fig. 1]. Informationen hierzu könnte möglicherweise die »Bedienungsanleitung« enthalten haben, die offenbar gemeinsam mit der Synthesekonstruktion selbst nach St. Petersburg verschifft wurde. Von beidem fehlt heute jeder Hinweis auf ihren späteren Verbleib.

Hinzu kommt das ganz allgemeine Problem, dass KRATZENSTEIN sich einer sehr ungewöhnlichen Terminologie zur Bezeichnung der einzelnen Bauteile an seinen Vokalpfeifen bedient. Es finden einerseits Begriffe aus

<sup>99</sup>Der abgebildete Resonator für U (Tab. II, Fig. 14) entspricht dem Modell 3 für O bis auf geringfügige Modifikationen [Kratzenstein 1782, 379].

der Anatomie Verwendung wie *Larynx* und *Epiglottis* und andererseits Begriffe aus dem Orgelbau wie *Pfeife*, *Kopf*, *Nuss* etc. Letztere werden jedoch oftmals in einer anderen Bedeutung verwendet als im Orgelbau üblich. So bezeichnet der Begriff *Pfeife* bei KRATZENSTEIN beispielsweise den *Resonator*, während er im Orgelbau üblicherweise für den gesamten Pfeifenkörper steht. Mit dem Begriff *Larynx* bezeichnet er zumeist den Komplex aus *Kehle* und *Zungenblatt*, scheint ihn bisweilen aber auch für die gesamte *Pfeife* zu verwenden. Diese terminologische Vertauschung und Inkonsequenz erschwert das Verständnis seiner Konstruktionsbeschreibungen in ganz erheblichem Maße.

Eine Rekonstruktion der »Vokalorgel« wäre aufgrund der hier dargelegten Voraussetzungen folglich nur durch sehr zeit- und materialaufwendiges »trial and error« möglich, für das KRATZENSTEINS Angaben zwar die grundsätzliche Richtung vorgeben könnten, aber keine konkreten Bauanweisungen. Dies war jedoch im Rahmen dieser Arbeit weder hinsichtlich der materiellen noch der zeitlichen Kapazitäten darstellbar und musste daher als Folgeprojekt zurückgestellt werden.

Anders als KRATZENSTEINS Ansatz zur Vokalsynthese ist die Arbeit WOLFGANG VON KEMPELENS vordergründig außerordentlich umfassend dokumentiert. Gleichwohl bleiben im konstruktiven Detail sowohl eine Vielzahl von Unklarheiten als auch denkbaren Veränderungen der Sprachmaschine. Das folgende Kapitel befasst sich mit den für die Saarbrücker Replik unternommenen und projektierten Modifikationen.

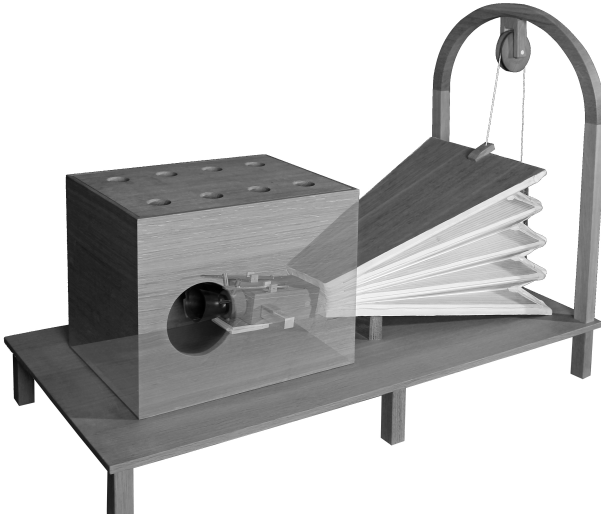
# Kapitel 8

## Kempelens Sprachmaschine

### 8.1 Die Saarbrücker Repliken

In den Jahren 2007–2009 wurden vom Autor insgesamt drei Repliken der Sprachmaschine WOLFGANG VON KEMPELENS angefertigt. Maßgeblich hierfür waren zunächst KEMPELENS eigene Angaben im »*Mechanismus*«, die jedoch für zahlreiche zentrale Details auf experimenteller Basis ergänzt werden mussten (vgl. [Brackhane 2011, 108 ff.]). Der 2007 unter Mithilfe von DOMINIK BAUER und der HUGO MAYER ORGELBAU GMBH (Heusweiler/Saar) fertiggestellte Prototyp konnte im selben Jahr im Rahmen des 16<sup>th</sup> *International Congress of Phonetic Sciences* (ICPhS) in Saarbrücken vorgestellt werden. Im Jahre 2009 wurden dann auf Grundlage der mit diesem Prototyp gewonnenen Erfahrungen zwei baugleiche Repliken hergestellt, die sich heute am Lehrstuhl für Kommunikationsakustik der TU Dresden und im Heinz Nixdorf MuseumsForum Paderborn befinden (Abb. 8.1).

Im Rahmen seiner Masterarbeit führte der Autor mehrere Experimente zur Leistungsfähigkeit der Repliken durch [Brackhane 2011, 114 ff.]. Angeregt durch KEMPELENS eigene Erkenntnis, dass ein *Doppelrohrblatt* ein erheblich besseres Glottis-Analogon sei als eine *Lingualpfeife* [Kempelen 1791a, 392] und durch theoretische Überlegungen in der Sekundärliteratur, dass eine *durchschlagende Zungenpfeife* der Bauweise KRATZENSTEINS zu einem merklich besseren Ergebnis bei der Stimmqualität von KEMPELENS Sprachmaschine hätte führen müssen [Willis 1832, 402], wurde in diesem Kontext auch eine Versuchsreihe begonnen, die verschiedene alternative Anregungskonzepte auf ihre Funktionsfähigkeit überprüfen sollte. Im Folgenden sollen mehrere Versuchskonzepte vorgestellt werden,



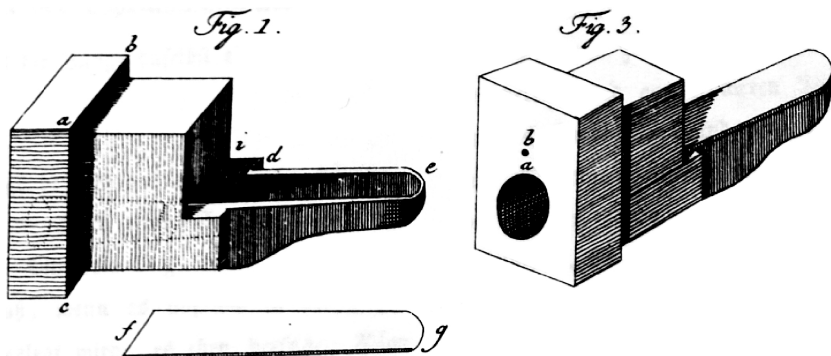
**Abbildung 8.1:** Die 2009er Version der Sprachmaschinen-Replik des Autors (Fotomontage, © Jan Braun/Heinz Nixdorf MuseumsForum).

mit denen die Veränderung der Stimmqualität unter Einsatz verschiedener alternativer Anregungskonzepte erforscht werden kann.

## 8.2 Alternative Anregungskonzepte

### 8.2.1 Kempelens aufschlagende Lingualpfeife

KEMPELENS Anregungskonzept erscheint aus heutiger Sicht verwunderlich. Hatte er im Laufe des »*Mechanismus*« noch explizit vom »betäubenden Geton« der *Vox humana* gesprochen [Kempelen 1791a, 390 f.] und mit der *membranösen Pfeife* und dem *Doppelrohrblatt* einer Oboe bzw. eines Dudelsacks gleich zwei zumindest theoretisch deutlich vorbildgetreuere Anregungskonzepte geschildert [Kempelen 1791a, 86 ff. und 390 ff.], so wählte er letzten Endes doch eine herkömmliche *aufschlagende Lingualpfeife*, von der er *Kopf* und *Kehle* unverändert übernahm (Abb. 8.2). Für *Zungenblatt* und *Resonator* entwickelte er zwar eigene Varianten, die jedoch ihr direktes Vorbild im Orgelbau nicht leugnen können. Was ihn letztlich dazu veranlasste, als Material der *Zunge* Elfenbein zu verwenden, muss offen bleiben. Vermutlich erhoffte er sich von diesem organischen Material einen weicheren Ton. Dadurch, dass er sowohl die Oberseite der *Kehle* als auch



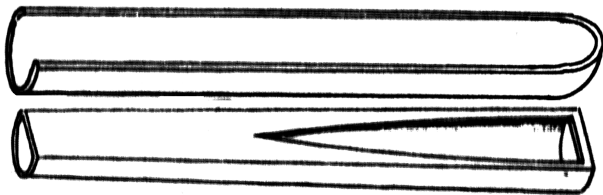
**Abbildung 8.2:** KEMPELENS Zungenpfeife (nach Kempelen (1791a), Tab. XVIII).

die Unterseite des *Zungenblattes* belederte (ersteres ist im Orgelbau üblich, letzteres allerdings nicht), reduzierte er jedoch ohnehin die klangbeeinflussenden Eigenschaften des Zungenmaterials auf ein Minimum. Auch KRATZENSTEIN hatte explizit auf Elfenbein als mögliches Material für *Zungenblätter* hingewiesen [Kratzenstein 1781, 24.]<sup>100</sup>. Es handelt sich hierbei jedoch weder im Orgelbau noch im sonstigen Instrumentenbau um ein gebräuchliches Material für klingende Teile eines Musikinstruments.

### 8.2.2 Varianten der aufschlagenden Lingualpfeife

Für den Saarbrücker Prototyp der Sprachmaschine wurden vom Autor sukzessive vier verschiedene Konfigurationen einer *aufschlagenden Lingualpfeife* eingesetzt und deren klangliche Unterschiede im Rahmen seiner Magisterarbeit knapp angesprochen [Brackhane 2011, 109 ff.]. Zu Beginn der Konstruktionsarbeiten an der Replik war noch keine *Zungenpfeife* nach den Vorgaben aus KEMPELENS »*Mechanismus*« verfügbar. Diese musste eigens angefertigt werden, was zum damaligen Zeitpunkt aus mehreren Gründen noch nicht möglich war. Stattdessen wurde von der unterstützend mitarbeitenden Orgelbaufirma HUGO MAYER GMBH (Heusweiler/Saar) eine dort überzählige moderne *Zungenpfeife* zur Verfügung gestellt. Diese unterscheidet sich in drei Punkten merklich von KEMPELENS Vorgaben, wobei nur zwei davon im gegebenen Kontext wesentlich sind: Bei der *Kehle* handelte es sich nicht um eine *Französische Löffelkehle*, wie sie KEMPELEN ausweislich seiner Tab. XVIII verwendet hat (vgl. Abb. 8.2 und Abb. 8.3

<sup>100</sup>vgl. Appendix, S. LVII



**Abbildung 8.3:** Schematische Zeichnung einer Französischen Löffelkehle (oben) und einer Deutschen Spitzkehle (unten) [nach Bauer 1996, 96]

oben), sondern um eine deutlich anders proportionierte *Deutsche Spitzkehle* (vgl. Abb. 8.3 unten und Abb. 8.4 links). Das ursprüngliche *Zungenblatt* besaß dementsprechend nicht nur ebenfalls andere Proportionen als das Vorbild, es bestand auch – wie im Orgelbau üblich – aus Messingblech anstatt des von KEMPELEN vorgegebenen Elfenbeins.<sup>101</sup>

Nachdem es zu einem späteren Zeitpunkt möglich war, eine eng an KEMPELENS Vorgaben orientierte *Zungenpfeife* zu bauen (vgl. Abb. 8.5 links), wurde für diese zunächst auch ein *Zungenblatt* aus Messing angefertigt. Kurze Zeit später konnte der Autor mehrere alte Klaviertastenbeläge aus Elfenbein erhalten, aus denen sich durch ebenso behutsames wie langwieriges Schleifen akzeptable Zungenblätter für beide vorhandenen *Zungenpfeifen* herstellen ließen (Abb. 8.4 rechts). Lediglich das Elfenbeinzungenblatt für die vorbildgetreue *Zungenpfeife* wurde nach KEMPELENS Vorgaben an seiner Unterseite mit dünnem Rinder-Spaltleder beledert. Bei den übrigen drei *Blättern* erwies sich eine solche *Beledering* als zumindest der Funktionssicherheit nicht förderlich und wurde daher nicht dauerhaft umgesetzt.

Die vier *Zungenblätter* weisen – wiewohl sie nur zwei unterschiedliche schwingende Längen besitzen (vgl. Abb. 8.4 rechts) – unter kontrolliert identischen Bedingungen deutliche Unterschiede in ihrer Grundfrequenz auf (Tab. 8.1).<sup>102</sup> Auffällig ist hierbei, dass die beiden gegensätzlichsten *Blätter* – das breite Elfenbeinzungenblatt nach KEMPELEN und das schmale Messingzungenblatt – sehr ähnliche Frequenzwerte aufweisen. Es dürfte dies ein Resultat aus den Stärken der *Zungenblätter* sein, wenn-

<sup>101</sup>Die in Abb. 8.4 erkennbare erhebliche »passive« Länge des zugehörigen Metallblattes ist durch die Konstruktion des *Kopfes* bzw. der *Kehle* bedingt und hat keinen Einfluss auf den Klang der Pfeife. Bei der Anfertigung des analogen *Zungenblattes* aus Elfenbein wurden diese Proportionen übernommen.

<sup>102</sup>Grundsätzlich bewirkt bereits eine geringfügige Erhöhung des *Winddrucks* einen Anstieg der Grundfrequenz der Pfeife (und umgekehrt). Daher ist für verwertbare Ergebnisse ein identischer *Winddruck* zwingend notwendig.

gleich diese vordringlich für den Obertonaufbau verantwortlich gemacht wird [Adelung 1982, 111]. Insbesondere die Stärke der *Zungenblätter* aus Elfenbein war indessen aufgrund ihrer Herstellungsweise nicht exakt zu kontrollieren.

Die spektrale Zusammensetzung dieser beiden augenscheinlich so unterschiedlichen Zungenblätter weist ebenfalls einige Ähnlichkeiten auf. Insbesondere im Bereich der ersten Harmonischen besteht eine erhebliche Ähnlichkeit, teilweise sogar Übereinstimmung der Frequenzlagen. Die Intensität der jeweiligen Harmonischen unterscheidet sich weitaus deutlicher (Abb. 8.6).<sup>103</sup> Über die Syntheseergebnisse unter Verwendung der verschiedenen Zungenblätter wurde bereits in [Brackhane & Trouvain 2008, Brackhane 2009] berichtet.

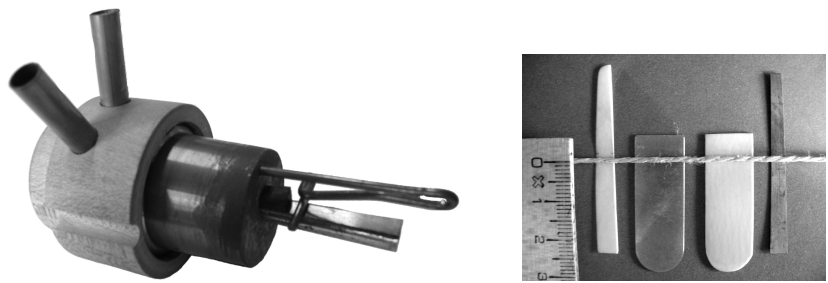
Eine weitere denkbare Modifikation betrifft die *Kehle*: Je weiter der Querschnitt einer *Kehle*, desto grundtöniger wird der durch sie erzeugte Ton, je flacher sie ist, desto schärfer und obertöniger wird er. Aus den hier beschriebenen Parametern Zungenbreite, Zungenstärke und Kehlenproportionen ergibt sich eine Vielzahl von möglichen Pfeifenkonfigurationen, wobei angenommen werden kann, dass bereits geringfügige Unterschiede eine merkliche Verbesserung oder Verschlechterung der synthetischen Stimme hervorrufen.

### 8.2.3 Durchschlagende Zungenpfeifen

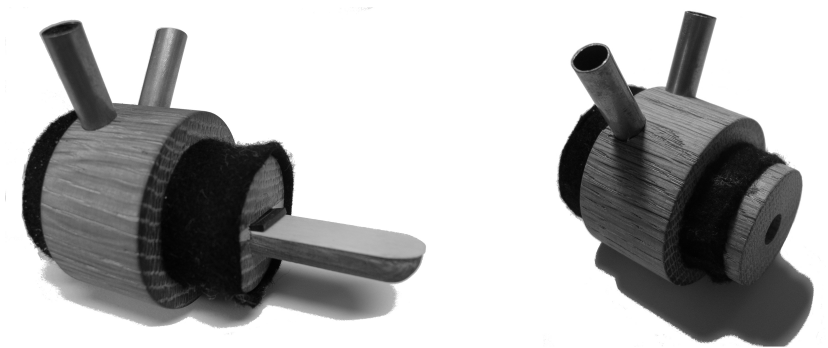
Aus der Retrospektive scheint die Sachlage zunächst eindeutig: CHRISTIAN GOTTLIEB KRATZENSTEIN hatte 1781 mit seiner preisgekrönten Schrift eine zwar bemerkenswerte, in ihrem anatomisch-artikulatorischen Kern aber unhaltbare Abhandlung über die Stimmgebung im Allgemeinen und das Wesen der fünf »Hauptvokale« A, E, I, O, U im Besonderen verfasst. Sein auf das *linguale Orgelregister* Vox humana rekurrerendes Konzept zur Synthese eben dieser Vokale führte zwar in dieser Disziplin in eine technologische Sackgasse, eröffnete aber dem Orgelbau eine gänzlich neue Form von *Lingualpfeifen*, die in der musikalischen Ästhetik des 19. Jahrhunderts eine enorme Bedeutung erhalten sollte. Gerade vor dem Hintergrund seiner übrigen wissenschaftlichen Leistung scheint die Episode »Vokalorgel« als (sprach-) wissenschaftsgeschichtliches Kuriosum für sich betrachtet ein knappes Unterkapitel wert, mehr aber wohl kaum.

Wie in Kapitel 3 bereits dargelegt wurde, steht KRATZENSTEINS Arbeit jedoch nicht als Solitär da, sondern ist zeitlich eingebettet in einen regelrechten Wettbewerb um die Entwicklung eines funktionstüchtigen Prototyps der Sprachsynthese. Insbesondere die Forschungen KRATZENSTEINS

<sup>103</sup>Spektra jeweils über vier Perioden aus der statischen Mitte des Klangs.



**Abbildung 8.4:** Die moderne Lingualpfeife für die Saarbrücker Sprachmaschinen-Replik (links) sowie die in dieser Replik verwendeten vier verschiedenen Zungenblätter (rechts) Die Linie bezeichnet die jeweils schwingende (=effektive) Länge der Blätter. Maße in cm.

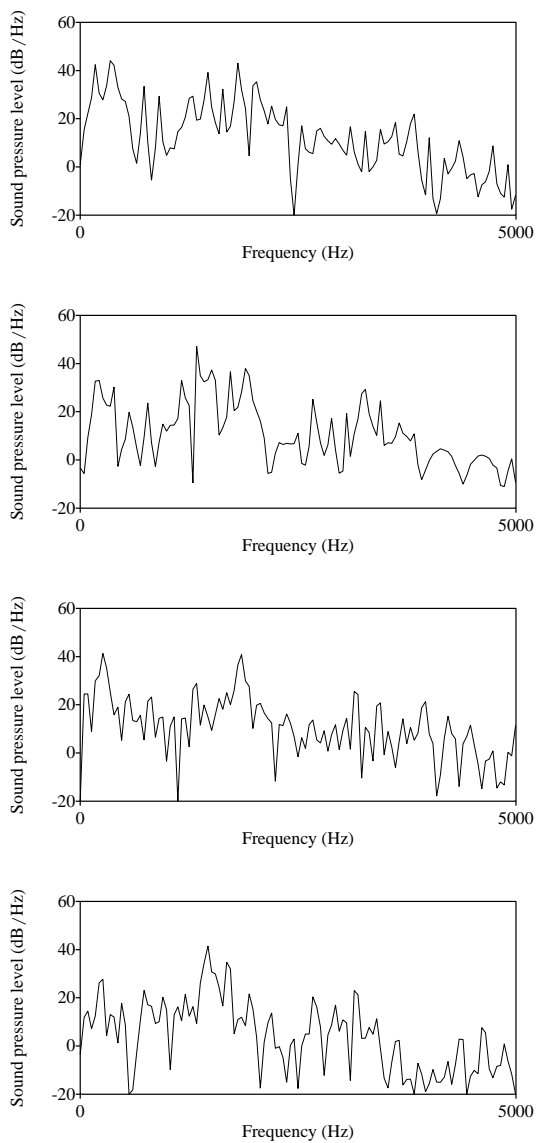


**Abbildung 8.5:** Köpfe der Saarbrücker Sprachmaschinen-Replik: Nach Kempelens Beschreibung (links) und ohne Kehle, für den Einsatz alternativer Anregungskonzepte (rechts).

**Tabelle 8.1:** Grundfrequenzen und  $F_1$ - $F_3$  (in Hz) der vier Zungenblätter der Saarbrücker SPRACHMASCHINEN-Replik. Die Messungen erfolgten jeweils bei vollständig zusammengesetzter Maschine.

Zungenblatt	$F_0$	$F_1$	$F_2$	$F_3$
Elfenbein schmal	243	1361	1708	2638
Messing breit	264	1132	1842	3227
Elfenbein breit	182	401	1773	2051
Messing schmal	192	1352	1877	3156





**Abbildung 8.6:** *Spektrale Zusammensetzung der Töne von breiter Elfenbeinzunge, schmaler Messingzunge, breiter Messingzunge und schmaler Elfenbeinzunge (v. o. n. u.).*

und KEMPELENS wurden hier in der Retrospektive immer wieder einander gegenüber gestellt, wobei Letzterer stets zuerkannt bekam, mit seiner Arbeit der Lösung des Sprachsyntheseproblems zumindest graduell näher gekommen zu sein. Es wurde gezeigt, dass KEMPELENS letztendliches Scheitern insbesondere auf die technologische Begrenztheit seiner Zeit zurückzuführen ist (vgl. [Brackhane 2011]). In manchen Einzelaspekten will es aber aus heutiger Sicht so scheinen, als ob zumindest eine Verbesserung auch schon für KEMPELEN hätte möglich sein müssen. Dies bezieht sich vor allem auf die zentrale Frage nach dem Anregungskonzept und damit der Stimmqualität.

KRATZENSTEIN hingegen wählte – wiewohl veranlasst durch völlig abwegige Ansichten zu Physiologie und Stimmgebung – eine in Europa vollkommen neue, weitestgehend unbekannte Variante der *Zungenpfeifen* und entwickelte sie (mit Hilfe eines Orgelbauers?) so weit, dass einer Verwendung in »richtigen« Orgelpfeifen schon bald nichts mehr im Wege stand. Die *durchschlagenden Zungenpfeifen* waren erfunden (bzw. neu entdeckt) und konnten ihren beispiellosen Siegeszug im Orgelbau des gesamten 19. Jahrhunderts antreten [Ahrens 2002, 44 f.; Ahrens 2003, 618].

Auch WOLFGANG VON KEMPELEN scheiterte mit seiner (zumindest in der Konzeption universellen) Sprachsynthese letzten Endes, nicht zuletzt auch an den technologischen Beschränkungen seiner Zeit. Ihm kann jedoch zugestanden werden, einen sowohl physiologisch als auch konzeptuell wesentlich überzeugenderen Ansatz gewählt zu haben, der zumindest auch auf seinen sehr fortschrittlichen Ansichten zur Physiologie und Artikulation beruhte. Sein »*Mechanismus der menschlichen Sprache*« stellt unstreitig einen Meilenstein innerhalb der frühen phonetischen Literatur dar.

Bei vergleichender Betrachtung dieser beiden so unterschiedlichen Synthesekonzepte drängt sich die Frage auf, ob nicht möglicherweise eine Kombination der Sprachmaschine KEMPELENS mit dem Anregungskonzept KRATZENSTEINS zu einer deutlich höheren Synthesequalität führen könnte. Warum hat KEMPELEN, wiewohl er einerseits im »*Mechanismus*« die ungenügenden Klangeigenschaften einer herkömmlichen *Zungenpfeife* ausdrücklich bemängelte und andererseits nachweislich KRATZENSTEINS Arbeit und somit auch dessen Anregungskonzept kannte, dieses nicht als Alternative begriffen? Aus heutiger Sicht erscheint es sicher, dass KEMPELENS Synthese unter Einsatz des kratzensteinischen Anregungskonzeptes zwar nicht »beredter« geworden wäre, die Stimmqualität vermutlich aber eine erhebliche Verbesserung erfahren hätte.

Auch die Versuche ROBERT WILLIS' legen einen derartigen Schluss nahe. Er schreibt, durchaus in Übereinstimmung mit KEMPELENS Äußerungen:

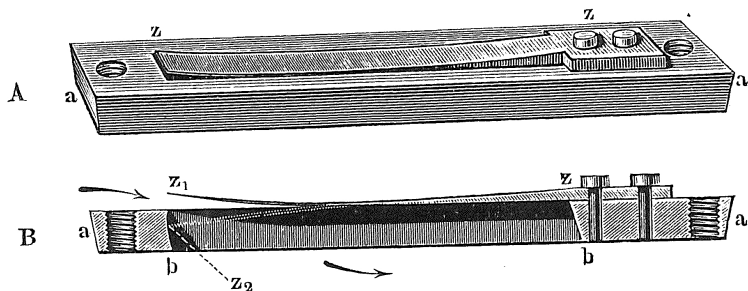
»Für das Gelingen dieser Versuche ist es wichtig, daß der Ton der Zungenpfeife so weich und rein wie möglich sey. Der rauhe Ton einer gewöhnlichen Rohrpfeife [*Rohrwerk* altertümlich für *Zungenpfeife*] ist zu diesem Zweck ganz untauglich, und daher sehen wir sowohl Kempelen als [auch] Kratzenstein bemüht, dieß Instrument zu verbessern. [...] Besser erreichte Kratzenstein seinen Zweck. Statt die Zunge auf den Rand der Röhre [Kehle] schlagen zu lassen, ließ er sie nämlich die Oeffnung nur genau verschließen, so daß sie eben Raum genug zur freien Vollbringung ihrer Schwingung erhielt. Dieß war eine sehr wichtige Verbesserung; denn bei sorgfältiger Ausführung dieser Construction erlangte der Ton der Zungenpfeife einen ganz neuen Charakter, und ward der menschlichen Stimme ähnlicher, als der irgend eines bis dahin bekannten Instruments; auch bekam er die nützliche Eigenschaft, daß er, innerhalb gewisser Gränzen, durch stärkeres Anblasen sich verstärken ließ, ohne seine Höhe zu ändern.« [Willis 1832, 402 f.]

Gemäß dieser Erkenntnis verwendete WILLIS zu allen seinen Versuchen *durchschlagende Zungenpfeifen* nach KRATZENSTEINS Vorbild [Willis 1832, 403]. Doch selbst er scheint trotz seiner prinzipiellen Erkenntnisse, dass eine Kombination von KEMPELENS Sprachmaschinenkonstruktion mit KRATZENSTEINS Anregungskonzept die Synthesequalität erheblich verbessert hätte, dieses nie experimentell überprüft zu haben.

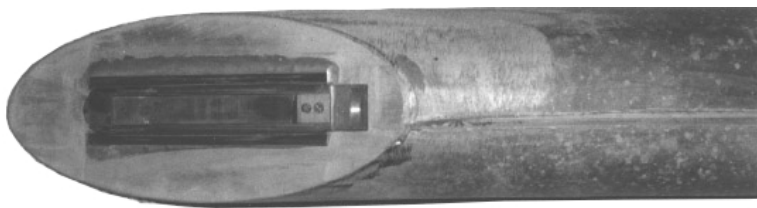
So bleiben bis zum heutigen Tage zwei experimentalphonetische desiderata: Einerseits eine Rekonstruktion der »Vokalorgel« KRATZENSTEINS und andererseits eben jene Modifizierung der Sprachmaschine KEMPELENS durch den Ersatz der herkömmlichen *Zungenpfeife* durch eine *durchschlagende*.

Die prinzipielle Möglichkeit zu Letzterem besteht, und dies sogar mit einem oberflächlich betrachtet geringem Aufwand: Es würde theoretisch völlig genügen, bei einer der existierenden Sprachmaschinen-Repliken die nach KEMPELENS detaillierten Vorgaben gefertigten *Kehle* durch eine solche mit geringfügig modifizierten Abmessungen auszutauschen, die für ein ebenfalls neu zu fertigendes *durchschlagendes Zungenblatt* geeignet ist. Praktisch stellt sich die Lage indessen deutlich schwieriger dar: Die Anfertigung einer derartigen *Kehle* erfordert ein ungleich präziseres Arbeiten, als es bei einer gewöhnlichen Kehle für eine *aufschlagende Zungenpfeife*

Fig. 28.



**Abbildung 8.7:** Ansicht und Schnittzeichnung einer durchschlagenden Harmoniumzunge nach dem Saugwindprinzip [?, 159].



**Abbildung 8.8:** Durchschlagende Zungenpfeife mit Harmoniumzunge.

der Fall ist. Dies liegt daran, dass die Öffnung der *Kehle* auf den Millimeterbruchteil genau so bemessen sein muss, dass das *Zungenblatt* einerseits, ohne zu verkanten darin frei schwingen kann, andererseits zwischen dem Rand des *Zungenblattes* und demjenigen der *Kehle* möglichst keinerlei Freiraum verbleiben darf, durch den Luft ungenutzt hindurch strömen könnte. Nicht zuletzt aufgrund dieses großen fertigungstechnischen Aufwandes werden *durchschlagende Zungenpfeifen* heute nur noch in Einzelfällen gebaut, ihre Fertigung ist dementsprechend kostspielig. Anders als bei *aufschlagenden Zungenpfeifen*, die jedenfalls theoretisch auch durch einen schreinerisch halbwegs begabten Laien gebaut werden können, ist dies im Falle der *durchschlagenden* Bauform ausgeschlossen.

Es besteht jedoch noch eine andere, weit kostengünstigere Möglichkeit, eine *durchschlagende Zungenpfeife* zu erhalten, die auch in der Hochzeit dieser Register im Orgelbau des 19. Jahrhunderts bisweilen Anwendung fand: Anstelle einer Konstruktion aus *Kehle* und *Zungenblatt* findet hier eine *Zungenpfeife* aus einem *Harmonium* der *Saugwind-Bauform*, wie auf

Abb. 8.7 und 8.8 zu sehen, Verwendung. Anstelle des Komplexes aus *Kopf* und der daraus herausragenden *Kehle* mit dem *Zungenblatt* ist bei Abb. 8.8 eine Harmoniumzunge auf der ansonsten verschlossenen Unterseite des *Resonators* so angebracht, dass der von unten durch den *Stiefel* strömende Wind das *Zungenblatt* in Schwingung versetzen muss, um in den *Resonator* zu gelangen.

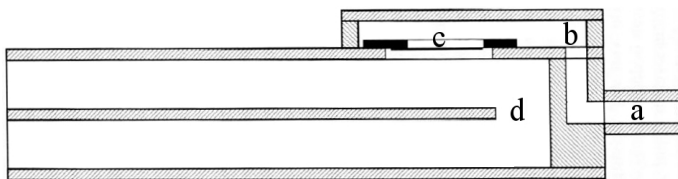
Eine derartige Konstruktion ist zwar ebenfalls nur mit einem gewissen Aufwand zu realisieren, doch sind Harmoniumzungen als Ausgangsmaterial auch heute noch verhältnismäßig günstig zu erwerben, ganz im Gegensatz zu den im Orgelbau verwendeten »richtigen« *durchschlagenden Zungenpfeifen*.

Eine dritte Variante wurde vom Autor im Rahmen seiner Magisterarbeit erprobt: Hierfür wurde eine herkömmliche Harmoniumzunge der *Saugwind*-Bauart in eine eigens angefertigte *Kanzelle* eingesetzt, die ihrerseits an einen *Kopf* angeschlossen wurde, der bis auf die fehlende *Kehle* baugleich mit der in der Saarbrücker Kempelen-Replik war (Abb. 8.5 rechts). Trotz zahlreicher Versuche unter Konsultation eines Orgelbauers gelang es jedoch nicht, eine mit der verwendeten *aufschlagenden Zunge* in ihren Abmessungen vergleichbare *durchschlagende Zunge* mit dieser Konstruktion befriedigend zum Klingen zu bringen.<sup>104</sup>

## 8.2.4 Zachariaszungen

Bei der sog. *Zachariaszunge* (Abb. 8.9) handelt es sich um eine sehr junge Konstruktionsvariante der *durchschlagenden Lingualpfeife*, die vom Trossinger Akustiker ERNST ZACHARIAS (\*1924) für den Musikinstrumentenhersteller HOHNER entwickelt wurde. Auch hier tritt die Luft zwar durch den sog. *Fuß* (a) ein, wird aber nicht wie sonst üblich direkt von diesem von unten, sondern durch einen sog. *Frosch* (b) seitlich in die Pfeife geleitet. Dort ist zwischen dem *Resonator* (d) und dem *Frosch* die eigentliche *Zunge* (c) befestigt. Gegenüber herkömmlichen *durchschlagenden Zungen* ist hier das *Zungenblatt* jedoch »gewendet« und wird beim Durchströmen der Luft zunächst aus der *Stimmlatte* herausgedrückt (vgl. auch Abb. 8.7). Hiermit würde normalerweise verhindert, dass das *Zungenblatt* überhaupt in Schwingung versetzt werden kann, eine Tonbildung somit unmöglich.

<sup>104</sup>Nach Abschluss dieser Arbeit konnte in Dezember 2014 schließlich doch noch eine durchschlagende Zungenpfeife in der hier skizzierten Weise für die Saarbrücker Sprachmaschinen-Replik hergestellt werden. Ihr Klang weist zwar erwartungsgemäß keinerlei Geräuschanteile auf, ist allerdings auch vollständig monoton, da sich Luftdruckveränderungen bei dieser Pfeifenbauform nur auf die Intensität und nicht auf die Frequenz des Klanges auswirken.



**Abbildung 8.9:** Schematische Schnittzeichnung einer Zachariaszunge [nach Eberlein 2011; 396].

Jedoch ändert sich dieser Umstand beim Anfügen eines *Resonators* (d) an die Konstruktion. Dieser zwingt dem *Zungenblatt* seine Eigenresonanz auf, da dieses »verkehrtherum« montiert eine optimale Dämpfung besitzt. Die reflektierte Stoßwelle des Resonators bewirkt mit ihrer Rückstellkraft schließlich das Schwingen der *Zunge*. Klangfarbe und (in gewissem Rahmen auch die Tonhöhe) werden allein durch den *Resonator* bestimmt. Genau wie bei *durchschlagenden Zungen* wirkt sich eine Erhöhung des Anblasedrucks lediglich auf die Intensität, nicht aber auf die Frequenz aus [Rohlf 1999].

Der Klang einer solchen – bislang sehr selten gebauten – *Zungenpfeife* wird von Musikern als vollkommen eigenständig beschrieben, weder mit *aufschlagenden* noch mit *durchschlagenden Zungenregistern* vergleichbar. Am ehesten trifft offenbar das Schlagwort eines »Horncharakters« zu: Ein »weicher«, grundtöniger, leicht verhaltener Klang, der jedoch stark von der Gestalt des *Resonators* abhängig ist.

Im Zusammenhang mit der Forschungsfrage, ob eine *durchschlagende Zungenpfeife* eine überzeugendere Synthesequalität in einer KEMPELEN-REPLIK gestatten würde, wäre auch der Einsatz einer solchen Zacharias-Zunge interessant gewesen. Es war jedoch nicht möglich, im gegebenen Zeitrahmen und mit den zur Verfügung stehenden limitierten Mitteln eine solche Zungenkonstruktion, die eigens von einem der sehr wenigen mit dieser Bauart bereits vertrauten Orgelbauer hätte angefertigt werden müssen, zu beschaffen.

### 8.2.5 Doppelrohrblätter

Wie bereits von KEMPELEN selbst beschrieben, handelt es sich bei einem *Doppelrohrblatt* um ein deutlich schlüssigeres Analogon zur menschlichen Stimmgebung als es bei einer *Zungenpfeife* – gleich, ob *auf-* oder *durchschlagender* Bauart – der Fall sein kann [Kempelen 1791a, 80 und 390 ff.]. Beim *Doppelrohrblatt* handelt es sich um ein aus zwei übereinandergelegten



**Abbildung 8.10:** Schematische Schnitzzeichnung (links, Gorski 2013) und Fotografie eines Doppelrohrblattes (rechts).

identischen schmalrechteckigen Stücken Pfahl- oder Schilfrohr bestehendes Mundstück, das als Anregung für eine ganze Reihe von klassischen Musikinstrumenten und für solche aus der Folklore stammenden (wie z. B. Oboe, Fagott, Dudelsack etc.) verwendet wird. Ähnlich wie bei den menschlichen Stimmlippen entsteht die Anregung der Luft hier durch die Schwingungen der beiden Elemente gegeneinander (Abb. 8.10).

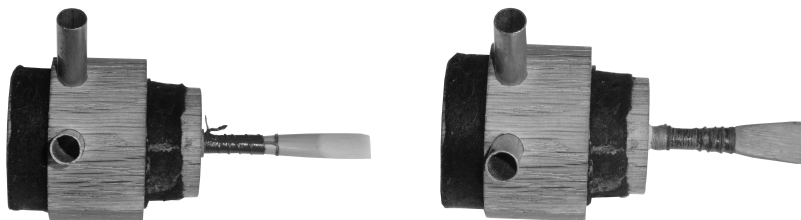
KEMPELEN war nach eigenem Bekunden auf das *Doppelrohrblatt* als Anregungskonzept gestoßen, indem er dessen Klang bei einer zufälligen Wahrnehmung für den eines singenden Kindes hielt [Kempelen 1791a, 391 f.], und benutzte es auch für seine ersten Syntheseveruche. Recht bald aber wandte er sich dann doch den zunächst geschmähten *Zungenpfeifen* zu [Kempelen 1791a, 399 ff.]. Die Gründe hierfür sind unbekannt, KEMPELEN selbst äußerte sich nicht zu seinem Sinneswandel.

Ähnlich wie im Falle der *durchschlagenden Zunge* stand auch hier die Frage im Raum, ob eine Ausstattung der KEMPELENSCHEN Sprachmaschine mit einem *Doppelrohrblatt* anstelle der *aufschlagenden Zungenpfeife* nicht für ein noch besseres Syntheseergebnis sorgen könnte. Entsprechende Versuche wurden vom Autor im Rahmen seiner Magisterarbeit konzipiert, konnten jedoch damals nicht durchgeführt werden [Brackhane 2011, 126].

Im Laufe mehrerer Gespräche mit einem Spieler von Doppelrohrblatt-Instrumenten wurde deutlich, um was für ein hochkomplexes und sensibles System es sich bei einem *Doppelrohrblatt* handelt.<sup>105</sup> Auch heute noch ist die Zusammenstellung der einzelnen Bauteile zu einem Mundstück und besonders die gebrauchsfertige Zurichtung des eigentlichen *Rohrblattes* eine diffizile Angelegenheit, für deren zahlreiche Arbeitsschritte es erheblicher Erfahrung und Übung bedarf. Vor diesem Hintergrund erst wird eine durch KEMPELEN berichtete Anekdote recht begreiflich. Es war ihm selbst gegen Geld nicht gelungen, von einem Dudelsackspieler eines seiner *Doppelrohrblätter* zu erhalten [Kempelen 1791a, 391 f.].

Bei der Suche nach einem passenden *Doppelrohrblatt* stellte sich heraus, dass das Mundstück eines Englischhorns (Alt-Oboe) zufällig exakt

<sup>105</sup>Dr. Wilfried Schütte (Mannheim) gebührt hier mein herzlicher Dank für seine Hilfsbereitschaft und intensive Beratung sowie die Überlassung mehrerer Doppelrohrblätter.



**Abbildung 8.11:** Die verwendeten Doppelrohrblätter im Adapter: Englischhorn-Blatt (links), Dudelsack-Spielpfeifen-Blatt (rechts).

in den Adapter für alternative Anregungsmechanismen der Saarbrücker Sprachmaschinen-Replik (Abb. 8.5 rechts) passte. Zusätzlich sollte das Doppelrohrblatt aus der *Spielpfeife* eines Dudelsacks in die Versuche mit einbezogen werden. Dies zum einen, da KEMPELEN ja explizit über seine Versuche mit einem von einem Dudelsack stammenden Doppelrohrblatt spricht. Zum anderen werden diese, anders als die Rohrblätter von Oboen, Fagotten etc. zum Spielen nicht direkt in den Mund genommen, sondern sind von einer *Windkapsel* umgeben, in die der Spieler Luft hineinbläst. Dieses Prinzip entspricht noch etwas mehr den baulichen Gegebenheiten von *Kempele's* Sprachmaschine als das gewöhnlicher Doppelrohrblätter. Das Doppelrohrblatt des Dudelsacks weist neben deutlich anderen Proportionen auch eine erheblich gröbere Fertigung auf, so verfügt es über keine sog. *Bahn*, sondern besitzt eine gleichbleibende Stärke.

Versuche mit diesen Blättern zeigten, dass im Vergleich zu einer *Zungenpfeife* ein erheblich höherer Luftdruck vonnöten ist, um ein Doppelrohrblatt zum Vibrieren zu bringen. Im konkreten Fall gelang es nur mit größter Mühe, mit dem Blasebalg der Saarbrücker Sprachmaschinen-Replik einen Luftdruck zu erzeugen, der die beiden Blätter zumindest schwach vibrieren ließ.

Als Einschränkung der Verwendbarkeit zeigte sich zudem, dass das Englischhorn-Blatt nur in angefeuchtetem Zustand (so, wie es auch beim Spiel mit seinem Instrument eingesetzt wird) verwendbar ist. Da hierfür jedes Mal die Demontage des gesamten Mundkomplexes notwendig wäre, schränkt dies die Verwendbarkeit von solchen Doppelrohrblättern merklich ein. Zuletzt war trotz der sehr schwachen Tonbildung erkennbar, dass der von beiden Blättern erzeugte Ton mit über 800 Hz erheblich zu hoch war, um als adäquates Stimmanalogon dienen zu können.<sup>106</sup>

<sup>106</sup>Die effektive Tonhöhe des mit einem solchen Doppelrohrblatt gespielten Instrumentes wird vom Instrumentenkörper bestimmt.



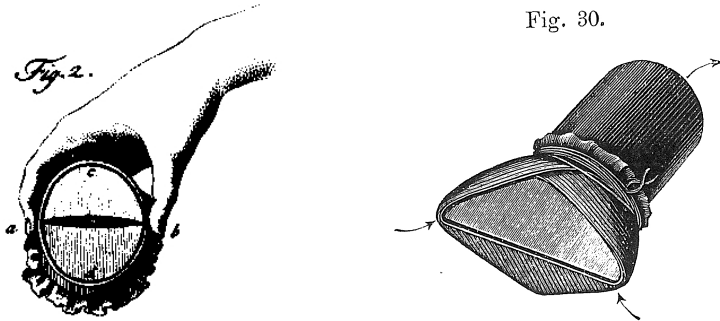
Die Verwendung eines Fagott- oder gar Kontrafagott-Blattes könnte möglicherweise einen Ton hervorbringen, der für eine Kinder- oder Jugendlischenstimme eine adäquate Tonhöhe aufweist. Es wurde aber im Rahmen der hier vorgestellten Versuche überdeutlich, warum KEMPELEN vermutlich sehr schnell von der aus anatomischer Sicht naheliegenden Idee wieder abkam, Doppelrohrblätter als Anregungsinstrument einzusetzen: Die Anfertigung von Doppelrohrblättern ist eine heikle Angelegenheit, die auch von geübten Personen viel Geduld und Fingerfertigkeit erfordert. Zudem ist der erforderliche Luftdruck in jedem Falle verhältnismäßig hoch, auch wenn er in gewissen Maße durch die Feinjustierung eines Blattes verringert werden kann. Nicht zuletzt die hohe mechanische Empfindlichkeit von Doppelrohrblättern gegenüber vergleichsweise robusten Zungenpfeifen dürfte schließlich den Ausschlag dafür gegeben haben, dass KEMPELEN sich lieber auf dieses Anregungskonzept einließ.

### 8.2.6 Eine membranöse Pfeife

Im »*Mechanismus*« stellt KEMPELEN seine Theorie über das Wesen der Glottis eingehend dar [Kempelen 1791a, 80 ff.]. Auch wenn er hierbei nicht in allen Details von der modernen Forschung bestätigt wird, so liegt er mit seinen empirisch gewonnenen Erkenntnissen doch wesentlich näher an heutigem Wissen als sein »Konkurrent« KRATZENSTEIN. Seine von DODART stark beeinflusste Theorie legt er in § 46 seines »*Mechanismus*« detailliert dar und schildert die Glottis als eine hälftig geteilte Membran, die im Ruhezustand straff über ein zylindrisches Rohr gespannt ist. Wird nun durch das andere Ende des Rohres Luft geblasen, beginnen die Schnittkanten der Membran zu schwingen und erzeugen ein Geräusch oder gar einen Ton [Kempelen 1791a, 86 ff.] (vgl. Abb. 8.12).

Nach ihm verfolgten u. a. auch FÉLIX SAVART (1791–1841) und HERMANN VON HELMHOLTZ (1821–1894) noch diesen Ansatz [Savart 1826; Helmholtz 1863, 160 ff. und Fig. 30], die aber ebenso wie KEMPELEN an der praktischen Umsetzung des theoretisch der menschlichen Glottis nahestehenden Anregungskonzeptes scheiterten. SAVART beschreibt den Klang einer membranösen Pfeife wie folgt:

»The sounds of membranous pipes partakes of the quality of those of a flute, and from free mouth-pieces [Lingualpfeifen mit durchschlagenden Zungen]. They have no analogy with those of any musical instrument, and such pipes are in some respect the reverse of stringed-instruments.« [Savart 1826, 202]



**Abbildung 8.12:** *Membranöse Pfeifen bei KEMPELEN (links) [Kempelen 1791a, Tab. II] und HELMHOLTZ (rechts) [Helmholtz 1863, Fig. 30].*

Ähnlich wie im Falle der Zachariaszunge erwies sich auch insbesondere der zeitliche Aufwand einer Versuchsreihe mit einer Membranösen Pfeife im Rahmen dieser Arbeit als zu groß und musste daher zurückgestellt werden.

## 8.3 Perzeptionsexperimente mit Sprachmaschinen-Repliken

### 8.3.1 Ausgangslage

Über die Synthesequalität der Sprachmaschine KEMPELENS sind natürlicherweise keinerlei akustischen Zeugnisse überliefert. Hinzu kommt, dass anscheinend kein aus KEMPELENS Hand stammendes Exemplar erhalten geblieben ist. Lediglich eine Reihe von zeitgenössischen »Ohrenzeugenberichten« legt den Schluss nahe, dass diese Synthese auf eine wesentliche Anzahl Zuhörer den Eindruck einer recht natürlichen Kinderstimme gemacht haben muss [Hindeburg 1784, 51; Anonymus 1784a, 180; Anonymus 1784b, 483f; Ebert 1785, 47 ff.]. Auch KEMPELEN selbst attestierte seiner Maschine eine solche Stimmqualität, nicht ohne darauf zu verweisen, dass man bei einem Kind, das das Sprechen erst erlerne, mehr Fehler in seiner Aussprache akzeptiere als es bei einem erwachsenen Menschen der Fall sei [Kempelen 1791a, 442 f.]. Diese Charakteristik der Sprachmaschinenstimme als kindlich scheint vor allem im hohen Stimmtone begründet zu sein.

Im Rahmen eines Vortrages 2007 konnte der Autor bei der ersten öffentlichen Vorführung seines prototypischen Sprachmaschinen-Nachbaues fest-

stellen, dass manche der anwesenden Kleinkinder auf die mit einer klagend-fallenden Intonation synthetisierten Äußerungen »Mama« und »Papa« sehr intuitiv reagierten und ihrerseits begannen, nach ihren Eltern zu rufen, wiewohl für ihn selbst der Charakter einer Kinderstimme nur entfernt getroffen war. Mehreren anderen Repliken wurde ebenfalls eine nur sehr begrenzte Ähnlichkeit zu menschlicher Stimme attestiert [Goethe 1893, 154; Köster 1978, 118 und Broeke 1983, 15].

### 8.3.2 Datengrundlage und Durchführung

Im Kontext seiner Magisterarbeit wurden vom Autor mehrere Perzeptionsexperimente durchgeführt, in denen überprüft werden sollte, in wie weit »naive« Hörer ebenfalls die synthetisierten Äußerungen verschiedener Sprachmaschinen-Repliken als von einem Menschen stammend akzeptieren können. Diese Experimente wurden bereits in [Brackhane & Trouvain 2008; Brackhane 2009] knapp beschrieben. Die verwendeten Stimuli sind dem Audio- und Video-Anhang dieser Arbeit beigelegt.

Für diese Versuchsreihe wurden mit dem Saarbrücker Nachbau Aufnahmen der typischen Äußerungen »Mama« und »Papa« gemacht und durch verfügbare Aufnahmen des Nachbaues aus Budapest (NIKŁÉCZY & OLASZY 2001/02) ergänzt. Da für den Saarbrücker Nachbau drei funktionstüchtige Konfigurationen der Zungenpfeife verfügbar waren (vgl. Kapitel 8.2.2, die schmale Elfenbeinzunge war nicht zuverlässig verwendbar), wurden Aufnahmen mit allen dreien gemacht. Hinzu kamen je ein »Papa«- und ein »Mama«-Stimulus der Budapester Replik.

In einem ersten Perzeptionsexperiment wurde jeweils ein »Papa«- und ein »Mama«-Stimulus beider Sprachmaschinen-Nachbauten in ein randomisiertes Set von Alltagsgeräuschen (z. B. Tiergeräusche, Wasserplätschern, Lachen) mit einer jeweiligen Dauer von ca. 2-4 Sekunden eingebettet. Dieses Set von Aufnahmen wurde 10 Versuchspersonen vorgespielt; in einem offenen Antwortformat sollten sie jeweils unmittelbar im Anschluss an einen Stimulus schriftlich beschreiben, was sie glaubten, gehört zu haben.

In einem weiteren Versuch wurden zu den o. g. zweimal vier verschiedenen Realisierungen (je drei »Mama«- und »Papa«-Realisierungen der Saarbrücker Replik plus je eine der Budapester Replik) zwei nun verfügbare zusätzliche Stimuli aus Budapest eingesetzt. Es handelte sich um je eine »Papa«- und eine »Mama«-Synthese, die dort mit einer neuen *Zungenpfeife* angefertigt worden waren. Diese zehn randomisierten Stimuli wurden 12 Versuchspersonen vorgespielt mit der Aufforderung, diese dahingehend zu bewerten, ob sie akzeptabel nach der Stimme eines Kleinkindes klingen.

Die Versuchspersonen wurden darüber informiert, dass sie ausschließlich synthetische Sprache zu hören bekommen würden. Jeder Stimulus wurde im Laufe des Experimentes sechs Mal präsentiert. Die Längen der Realisierungen wurden in ihrer Dauer einander angepasst, indem ggf. einzelne Perioden aus dem stabilen Mittelteil des Signals ausgeschnitten wurden. Die Bewertung der Stimuli sollte auf einer Skala von 1 (»sehr akzeptabel«) bis 6 (»gar nicht akzeptabel«) angegeben werden. Keine der Versuchspersonen hatte am ersten Experiment teilgenommen.

In einem dritten, unmittelbar auf dem zweiten aufbauenden Perzeptionsexperiment wurde neben den *Zungenblättern* ein weiterer Faktor untersucht. Bislang waren die verwendeten Aufnahmen der Saarbrücker Replik mit einer *Windlade* angefertigt worden, die nicht auf KEMPELENS Maßangaben beruhte, sondern etwas größer war. Dies geschah ursprünglich, um die Arbeiten im Inneren der *Windlade*, die häufig nach dem Prinzip des »trial and error« erfolgen mussten, einfacher zu gestalten. Als Resultat aus den Ergebnissen des zweiten Perzeptionsexperiments wurde nun eine neue *Windlade* angefertigt, die sich in Material und Abmessungen eng an KEMPELENS Vorgaben bzw. an der hinsichtlich gerade dieses zentralen Bauteils mutmaßlich sehr alten Replik im Deutschen Museum München orientierte (vgl. [Kempelen 1791a, 415] und Kapitel 3.4.2 auf S. 69). Mit dieser neuen sowie der bisherigen *Windlade* wurden nun »Mama«-Stimuli mit unterschiedlicher Prosodie realisiert, wobei nur das breite Elfenbeinzungenblatt und das schmale Messingzungenblatt zum Einsatz kamen, da die anderen beiden mit der neuen *Windlade* nicht adäquat zum Ansprechen gebracht werden konnten. Als prosodische Muster wurden steigende und fallende Intonation bei gleichlangen Silben sowie bei monotoner Intonation die Silbendauern lang-lang, kurz-lang und kurz-kurz gewählt, resultierend in 36 Stimuli. Diese Stimuli wurden acht Versuchspersonen in randomisierter Reihenfolge präsentiert, wobei jeder Stimulus fünf Mal wiederholt wurde. Als Bewertungsskala wurde diejenige aus dem zweiten Experiment verwendet. Die Versuchspersonen (von denen wiederum keine an den vorangehenden Experimenten teilgenommen hatte) wurden darüber aufgeklärt, dass sie ausschließlich synthetische Sprache hören würden.

### 8.3.3 Ergebnisse

Im ersten Experiment wurden die beiden Sprachmaschinen-Stimuli von jeweils vier der zehn Versuchspersonen als kindliche Äußerung beschrieben. Die Übrigen beschrieben die Stimuli als synthetisch. Die Charakterisierung unterschied sich hierbei allerdings je nachdem, um welchen der Stimuli es sich handelte. Während der »Mama«-Stimulus durchgehend als Stimme

einer Spielzeugpuppe beschrieben wurde, war das Urteil beim »Papa«-Stimulus deutlich vielfältiger. Jeweils zwei Versuchspersonen beschrieben ihn als Singstimme oder maschinelle Stimme. Die verbleibenden zwei waren sich unsicher, in wie weit es sich um eine Sprech- oder eine Singstimme handle.

Im zweiten Experiment schnitten die beiden »Papa«-Stimuli der Budapester Replik mit einer Bewertung zwischen »2« und »3« am besten ab, während die drei Saarbrücker Stimuli Bewertungen zwischen »3« und »4« erhielten. Insgesamt wurden alle fünf »Papa«-Stimuli als überzeugender bewertet als die »Mama«-Stimuli. Zudem fiel die Bewertung der »Papa«-Stimuli deutlich differenzierter aus (Abb. 8.13).

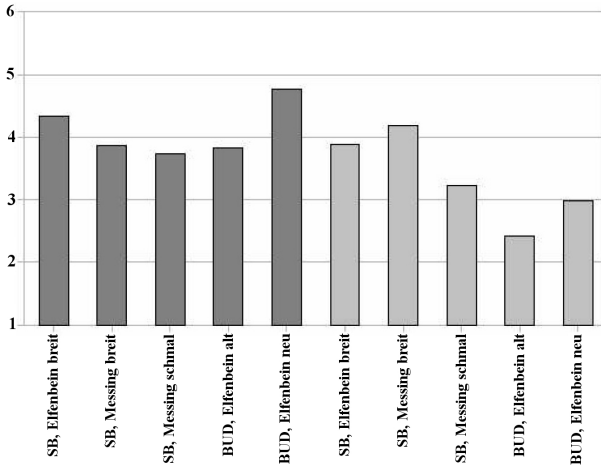
Die Wirkung mancher Stimuli war so stark, dass einzelne Versuchspersonen auch nach dem Test nicht glauben wollten, dass es sich nicht um eine menschliche Kinderstimme gehandelt habe.<sup>107</sup>

Sieben der neun am besten bewerteten Stimuli im dritten Experiment waren unter Verwendung der neueren, auf historischen Vorbildern basierenden Windlade sowie der ebenfalls auf historischen Vorbildern beruhenden *Zungenpfeife* mit breitem *Zungenblatt* aus Elfenbein entstanden. Unter den sieben am schlechtesten bewerteten Stimuli befanden sich ausschließlich solche, die mit der unhistorischen *Zungenpfeife* (jedoch mit beiden *Windladen*) produziert worden waren (Abb. 8.14). Die unterschiedlichen Dauerproportionen und Intonationsmuster hatten hingegen keinen Einfluss auf den Eindruck der Natürlichkeit.

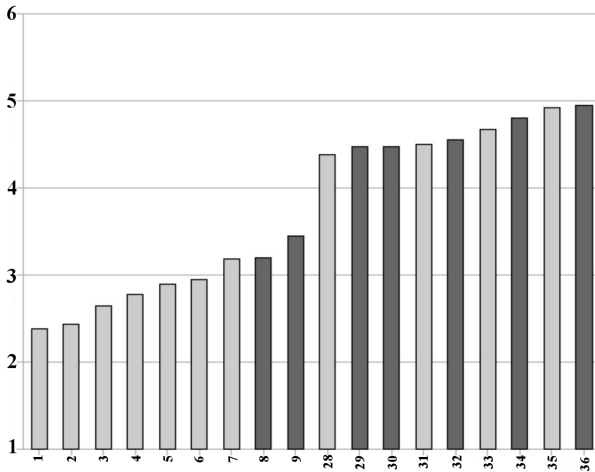
### 8.3.4 Diskussion

Die Experimente konnten zeigen, dass sowohl naive als auch nicht-naive Hörer die Produktionen einer Sprachmaschine nach KEMPELENS Vorbild durchaus als menschenähnlich empfinden können. Dies könnte unter anderem mit der Charakteristik der Synthese als »kindliche« Stimme zusammenhängen, hatte doch KEMPELEN selbst bereits festgestellt, dass man von einem Kind keine »lupenreine« Sprache erwarte [Kempelen 1791a, 442 f.]. Auffällig ist hierbei, dass die »Papa«-Stimuli im ersten Perzeptionsexperiment vielfältiger beschrieben wurden, sofern sie nicht als vom Menschen stammende Sprachäußerung akzeptiert wurden und zugleich im zweiten Experiment deutlich differenzierter bewertet wurden als die »Mama«-Stimuli. Freilich existiert zwischen den Stimuli der beiden Experimente nur eine kleine Schnittmenge, doch ist die Tendenz augenfällig.

<sup>107</sup> Um welche konkreten Stimuli es sich hierbei jeweils handelte, konnte aufgrund der randomisierten Präsentation nachträglich nicht mehr nachvollzogen werden.



**Abbildung 8.13:** Bewertungsraten der 10 »Mama«- (dunkelgrau) und »Papa«-Stimuli (hellgrau) des zweiten Perceptionsexperiments. »1« = »sehr akzeptabel«, »6« = »gar nicht akzeptabel«.



**Abbildung 8.14:** Bewertungsraten der neun am besten und neun am schlechtesten bewerteten Stimuli des dritten Perceptionsexperiments. Die hellgrauen Balken kennzeichnen Stimuli unter Verwendung der Windlade nach historischem Vorbild, die dunkelgrauen Stimuli unter Verwendung der früheren, unhistorischen Windlade. »1« = »sehr akzeptabel«, »6« = »gar nicht akzeptabel«.

Ob die unterschiedliche Bewertung der Stimuli im zweiten Experiment jeweils tatsächlich ausschließlich durch Konfigurierung der *Zungenpfeife* verursacht wurde, muss bezweifelt werden, da es mit der Sprachmaschine nicht möglich ist, Faktoren wie Intonationsverlauf und Artikulationsgeschwindigkeit vollkommen zuverlässig zu steuern. Da die neuere *Zungenpfeife* aus Budapest in beiden Fällen erheblich schlechter bewertet wurde als ihr älteres Pendant, kann allerdings vermutet werden, dass unabhängig von solchen Faktoren auch die konkrete Fertigung der *Zungenpfeife* eine Rolle spielt. In jedem Falle wäre es aber vor der Hintergrund der Versuchsergebnisse sehr wünschenswert, genauere Informationen über die Konstruktion der in der Budapester Replik verwendeten Zungenpfeife zu erhalten. Dies war dem Autor jedoch bislang trotz mehrerer Anfragen nicht möglich. Da – bis auf eine Ausnahme – im zweiten Experiment alle Stimuli mit schlechter als »3« bewertet wurden, muss der Schluss gezogen werden, dass auch bei solch rudimentären Synthesen das Ergebnis letztendlich nicht zu überzeugen vermochte. Ein Vergleich der Sprachmaschinen-Stimuli mit Synthesen verschiedener elektronischer Synthesen im Rahmen eines weiteren Perzeptionsexperimentes könnte hier möglicherweise helfen, die Qualität der Sprachmaschinen-Synthese zu bewerten.

Der deutliche Einfluss der zunächst eher nebensächlich erscheinenden *Windladen*-Abmessungen auf die Akzeptabilität der Synthese überrascht, unterschieden sich die beiden verwendeten Modelle nur um wenige Zentimeter voneinander. Leider gibt KEMPELEN keinerlei Gründe für die von ihm mitgeteilten Maße an. Es kann jedoch nicht ausgeschlossen werden, dass sich diese womöglich versuchsweise ermittelten Abmessungen für die verwendete *Zungengröße* als ideal erwiesen hat. Auch der scheinbar sehr geringe Einfluss der prosodischen Muster auf die Akzeptabilität der Stimuli ist bemerkenswert, wurden doch »Mama«-Realisierungen mit unnatürlich großer Silbenlänge nicht grundsätzlich schlechter bewertet als solche mit plausiblerer Prosodie. Eventuell kann auch dieses Phänomen damit erklärt werden, dass Zuhörer von einer kindlichen Stimme nicht zwingend prosodisch korrekte Sprachrealisierungen erwarten.





# Teil III

## Resumée



»The desire to imitate the human voice  
is as ancient as history  
and as pervasive as human culture.«  
[Hankins & Silverman 1995, 178]

Wenn auch dem im voranstehenden Zitat erwähnten Streben nach Lösungen zur synthetischen Nachahmung menschlicher Sprache im Rahmen dieser Arbeit nicht bis zu den Anfängen der Geschichtsschreibung nachgegangen werden konnte, so zeigt die hier ab dem siebzehnten Jahrhundert skizzierte Historie doch, wie gleichermaßen zielstrebig, aber auch unidirektional diese Aufgabe angegangen wurde: Bereits ATHANASIUS KIRCHER (1602-1680) beschrieb als Synthesekonzept die Nachbildung der physiologischen und artikulatorischen Gegebenheiten beim Menschen samt ihrer Funktionalität (so gut als man diese verstanden zu haben glaubte) unter gleichzeitiger Verletzung dieses Vorsatzes durch die Verwendung einer *aufschlagenden Lingualpfeife* als Glottisanalogon. Die Überzeugung, durch den Umweg über die Nachmodellierung des Vokaltrakts zu einer authentisch klingenden Sprachsynthese gelangen zu können, hält sich mit dem »Waseda Talker« von FUKUI et al. (2008) bis in die Gegenwart. Ähnlich langlebig zeigt sich auch der Ansatz der Verwendung von *Zungenpfeifen* für die Anregung, die zuletzt in den Konstruktionen von RIESZ (1903–1974) thematisiert wird.

Sowohl die Idee, die menschliche Anatomie nachzubilden als auch die Idee, eine *Zungenpfeife* als Glottisanalogon zu verwenden, sind leicht nachvollziehbar. Es musste gerade in Zeiten, als über die physikalischen Eigenschaften von Schall im Allgemeinen und von Sprachschall im Besonderen noch nichts bekannt war, zwangsläufig als das einzig sinnvolle Mittel zum Zweck erscheinen, die Artikulationsorgane des Menschen (bzw. was man dafür hielt) so gut es irgend ging mechanisch nachzubilden. Die engen Grenzen dieses Ansatzes zeigte indessen bereits KEMPELEN (1743-1804) in seinem »*Mechanismus*« von 1791 deutlich auf. Dennoch wurde in den folgenden zwei Jahrhunderten noch mehrfach versucht, auf diesem Wege zum Erfolg zu gelangen. Die Tatsache, dass heute über so faszinierende Konzepte wie beispielsweise die Vokaltrakte von RIESZ nahezu nichts mehr bekannt ist, lässt darauf schließen, dass die faktische Synthesequalität auch post-kempelenscher Synthesekonzepte erheblich zu wünschen übrig ließ. Dies gilt auch und insbesondere für die »Euphonia« FABERS (Lebensdaten unbekannt). Gerade über diese Synthesekonstruktion existieren zahlreiche beinahe überschwängliche zeitgenössische Berichte, die ihr eine wenn schon

nicht authentische, so doch zumindest beeindruckende Synthesequalität attestierten. Trotzdem verliert sich auch ihre Spur, ähnlich ihrem Vorbild, der Sprachmaschine KEMPELNS, alsbald im Dunkel der Zeit. Hier wie auch bei nahezu allen anderen beschriebenen mechanischen Sprachsynthesekonzepten stellt sich die Frage, wie dieser Widerspruch zwischen positiver zeitgenössischer Berichterstattung und dem raschen in Vergessenheit Geraten nach wenigen Jahrzehnten zu erklären ist.

Die Syntheseforschung CHRISTIAN GOTTLIEB KRATZENSTEINS (1723-1795) weist einige Ungereimtheiten auf. Bei näherer Betrachtung seiner insbesondere im 19. Jahrhundert heroisierten Leistung zeigt sich rasch, dass sein »*Tentamen*« den fraglichen Gegenstand weitaus weniger substantiell und detailliert behandelt als man annehmen sollte. Ungeachtet der vermutlichen Preisauslobung »ad personam« muss konstatiert werden, dass KRATZENSTEINS eigentliche Leistung nicht in der anatomisch-physiologischen Erklärung der Vokalartikulation lag, sondern in der Verwendung eines neuartigen Anregungskonzeptes für seine ansonsten höchst mediokre Vokalsynthese. Inwieweit gerade die Entwicklung der *durchschlagenden Zungenpfeife* tatsächlich KRATZENSTEIN und nicht zumindest in wesentlichen Teilen einem Orgelbauer (FRANZ KIRSCHNICK (1741-1802)?) zuzuschreiben ist, muss vorerst offen bleiben. Die Entwicklung der *durchschlagenden Zungenpfeife* stellt fraglos eine ingeniöse und handwerklich höchst anspruchsvolle Leistung dar, die trotz KRATZENSTEINS Attribuierung als hervorragender »Mechanicus« eher einem versierten Orgelbauer zuzutrauen wäre. Ebenso machen die vier *Resonator*-Formen der »Vokalorgel« misstrauisch, lassen sich doch zumindest derer drei im zeitgenössischen Orgelbau vielfach wiederfinden. Ihre Beschreibung wie die der gesamten »Vokalorgel« bleibt indessen sehr vage. Um so interessanter wäre es zu wissen, was in der »Betriebsanleitung« stand, die KRATZENSTEIN allem Anschein nach mit der Maschine selbst nach St. Petersburg schickte. Doch diese scheint entweder verloren zu sein oder unzugänglich in den Archiven der Russischen Akademie der Wissenschaften zu schlummern.

Das Phänomen der frühen und »alternativlosen« Festlegung auf *Zungenpfeifen* als Anregungskonzept kann wohl nur teilweise durch die partielle Ähnlichkeit solcher Orgelpfeifen mit der menschlichen Glottis erklärt werden. Wäre es hier ausschließlich um eine enges Analogon gegangen, hätte es mit den von KEMPELEN (und in der Folge u. a. auch von HELMHOLTZ) beschriebenen Alternativen *Doppelrohrblatt* und *membranöse Pfeife* deutlich schlüssigere Lösungen gegeben. Dass diesen jedoch dauerhaft *Lingualpfeifen* vorgezogen wurden, mag zum einen durch eine gewisse Bequemlichkeit erklärbar sein. Man besaß ein halbwegs stimmiges und annehmbar funktionierendes Konzept, mit dem man sich zunächst be-

gnügte und dessen Verbesserung man zugunsten der Forschung an anderen Aspekten der Sprachsynthese einstweilen zurückstellte. Zumindest bis ins neunzehnte Jahrhundert hinein scheint aber zum anderen auch die ebenso irrationale wie hartnäckige Idee einer authentischen *Vox humana* (VH) eine nicht zu unterschätzende Rolle gespielt zu haben, wobei es sich hierbei vermutlich hauptsächlich um eine weitenteils unbewusste Fortschreibung alter Legenden gehandelt haben dürfte.

Ähnlich wie im Falle der häufig positiven, bisweilen geradezu überschwänglichen Berichterstattung über die Synthesequalität verschiedener Sprachmaschinen des achtzehnten und neunzehnten Jahrhunderts dürften auch bei der ebenfalls sehr positiven gearteten Attribuierung mancher VH-*Register* autosuggestive Faktoren eine Rolle gespielt haben. Für die Sprachmaschine WOLFGANG VON KEMPELEN ist – im Gegensatz zu eigentlich allen anderen Sprachmaschinen – gut belegt, wie sich ihr Schöpfer durchaus manipulativ das Staunen des Publikums zu sichern wusste (ohne, dass es dabei mit offenkundigem Betrug zugegangen wäre). Auch das lebhaftes Presseecho dürfte von KEMPELEN zumindest teilweise selbst instrumentiert worden sein.<sup>108</sup> Für die zahlreichen als besonders überzeugend und authentisch beschriebenen VH-*Register* hingegen, von denen im Kapitel 2 einige Beispiele angeführt wurden, wissen wir nicht, auf welche Weise die jeweiligen Berichtersteller zu ihrem Urteil gelangten. Gerade im Falle der VH der GABLER-ORGEL zu Weingarten (1750) dürfte in der zunächst absurd erscheinenden Teufels-Legende ein wahrer Kern stecken, zeugt doch die Gestalt just dieser Pfeifen unzweifelhaft von dem hartnäckigen Ringen eines auch sonst die Perfektion bis an den Rand der Besessenheit treibenden Orgelbauers um ein *Register*, das seinem als programmatisch verstandenen Namen so gerecht wie möglich werden sollte.

In jedem Falle bildete sich im achtzehnten und – begünstigt durch die schwärmerische Grundhaltung der Romantik – auch noch im neunzehnten Jahrhundert ein regelrechtes Legendengeflecht um dieses *Register*, wie die im zweiten Kapitel angeführte Anekdote um dieses wohl berühmteste VH-*Register* überhaupt, das der Benediktinerabtei Weingarten, zeigt. WERNER LOTTERMOSER war der erste, der seit den 1930er Jahren im Rahmen seiner grundlegenden akustischen Untersuchungen zum Orgelklang versuchte, objektive Indizien für den vokalähnlichen Klang der VH zu finden. Wiewohl seine Ergebnisse aufgrund der seinerzeit noch recht

---

<sup>108</sup>Insbesondere die Berichte von HINDENBURG (1784) und EBERT (1785) sind wohl kaum ohne direkten Kontakt der Autoren zu KEMPELEN zustande gekommen. Für den chronologisch ersten Bericht von WINDISCH (1783) ist dieser Kontakt sogar durch den Autor selbst belegt. WINDISCH war offenbar eng mit KEMPELEN befreundet [Windisch 1783a, 3 f.].

begrenzten technischen Analysemöglichkeiten nur äußerst stichprobenartig ausfallen mussten, konnten seine Angaben als Ausgangspunkt für eigene Untersuchungen verwendet werden.

Die im Rahmen dieser Arbeit durchgeführten Analysen konnten einerseits LOTTERMOSERS Resultate nicht ohne weiteres bestätigen, gleichwohl lag es aufgrund der Ergebnisse nahe, den untersuchten vier *VH-Registern* hinsichtlich ihres Klangspektrums eine zumindest graduell größere Ähnlichkeit zu menschlichen Vokalen zuzusprechen, als es für die parallel untersuchten historischen *Lingualregister Cromorne* und *Trompette* der Fall sein konnte. Die ergänzend durchgeführten Perzeptionstests stützen diese Feststellung insofern, als zumindest peripheren Tönen des *Registers* von den Versuchspersonen sehr deutliche Vokalqualitäten zugeschrieben wurden.

Künftige Versuche sollen die hier vorgestellten Untersuchungen erweitern. Insbesondere soll Vergleichsmaterial von anderen Zungenpfeifen, die nicht dem Bauprinzip der VH unterliegen, und von *Labialpfeifen* verschiedener Bauarten erhoben werden. Mit diesen Aufnahmen wäre es sodann möglich, einen aussagekräftigeren Vergleich über die Vokalähnlichkeit von VH-Pfeifen anzustellen.

Doch nicht nur im Bereich der VH gilt es, die Forschungen künftig fortzusetzen und zu erweitern. Die Kapitel 7 und 8 haben versucht aufzuzeigen, wie viel an experimentalphonetischen Desideraten in diesem Bereich noch der Bearbeitung harret. Ähnlich wie die Fortsetzung der VH-Versuchsreihe werden auch diese Fragestellungen Stoff für die Forschungen der kommenden Jahre sein.

# Danksagung

Zuvörderst gilt es, meinem Doktorvater Prof. Dr. BERND MÖBIUS (Universität des Saarlandes, Saarbrücken) zu danken, bei dem ich nicht nur mit meinem Konzept auf offene Ohren stieß, sondern auch stets mit tatkräftiger Unterstützung rechnen konnte. Prof. em. Dr. WILLIAM J. BARRY (Universität des Saarlandes, Saarbrücken) erklärte sich nicht nur bereit, trotz seines Ruhestandes als Zweitgutachter dieser Arbeit zu fungieren, sondern hatte auch bereits mit seiner wohlwollenden Unterstützung der thematisch verwandten Magisterarbeit seinen Anteil daran, meine Begeisterung für diesen Themenkomplex zu fördern. Sodann habe ich Prof. Dr. RÜDIGER HOFFMANN (TU Dresden) zu danken, mit dem mich mittlerweile mehrere Jahre der Kempelen-Forschung verbinden. Er trug nicht nur einen Gutteil technische Fachkompetenz bei, sondern animierte mich auch, meinen Blick auf die Entwicklungen des 19. Jahrhunderts auszuweiten. Ein ganz besonderer Dank gilt meinem Mentor und »Anstifter« Dr. JÜRGEN TROUVAIN (Universität des Saarlandes, Saarbrücken), der mich im Jahre 2007 erstmals für die Thematik begeisterte und mir seither in großem Umfang Rat und Unterstützung zukommen ließ, nicht zuletzt durch zahlreiche anregende, den Blick schärfende Diskussionen. Ohne ihn wäre diese Arbeit nie zustande gekommen.

Es versteht sich von selbst, dass eine größere Arbeit – zumal wenn sie thematisch vielschichtig ist – schwerlich im Alleingang des Autors entstehen kann. Nicht nur, dass es stets der Korrektur und konstruktiven Kritik durch Dritte bedarf. Auch umfangreiche Recherchen sind fast nie möglich ohne die kooperative Zuarbeit anderer. Mir war es vergönnt, im Laufe meiner Arbeit in den verschiedenen Archiven und Institutionen beinahe ausschließlich auf Gegenüber zu treffen, die nicht nur meine Fragen bereitwillig beantworteten, sondern sich auch von meinem Wissensdrang anstecken ließen und bisweilen sogar von sich aus weiterrecherchierten. Auch zahlreichen Kollegen vom Institut für Deutsche Sprache (IDS), Mannheim,

habe ich herzlich für ihre so bereitwillige und mannigfaltige Unterstützung zu danken.

Zumindest die Namen dieser vielen Helfer, Ratgeber und Zuarbeiter seien hier genannt: Prof. Dr. Dr. h.c. OTTO BIBA (Gesellschaft der Musikfreunde, Wien), Dr. RAINER BRACKHANE (Herford), ERICH KAUFHOLD (Barth), Prof. Dr. ANDREAS KLEINERT (Universität Halle-Wittenberg), JACQUELINE KUBZAK (IDS, Mannheim), LOTHAR LAMBACHER (Kunstgewerbemuseum Berlin), Dr. ELISABETH LINK (IDS, Mannheim), Dr. RUTH MELL (IDS, Mannheim), Dr. JULIANE MIKOLETZKY (TU, Wien), Prof. Dr. JOHN OHALA (Berkeley), MONIKA POLSCHMIDT (IDS, Mannheim), Dr. STEPHAN PUILLE (Berlin), WOLFGANG RATHKE (IDS, Mannheim), Dr. SUSANNE REHN-TRAUBE (Deutsches Museum, München), MELANIE SCHMÄH (Worms), JOACHIM SCHREIBER (Simmern), Dr. WILFRIED SCHÜTTE (IDS, Mannheim), Prof. Dr. MARIJA SMIRNOVA (Lomonossow-Universität, Moskau), Dr. STEFAN STEIN (Nixdorf MuseumsForum, Paderborn), KARSTEN UNVERRICHT (Mannheim), Prof. Dr. ELISABETH VAUPEL (Deutsches Museum, München), Dr. ANNA VOLODINA (IDS, Mannheim).

Herzlich danken möchte ich auch all jenen Kirchengemeinden und Organisten, die ihre wertvollen historischen Orgeln samt VH nicht nur »anlässlich der sonntäglichen Orgelmusiken ihre Wirkung auf fromme Gemüter live und subjektiv erlebbar«, sondern auch meinen gänzlich säkularen Forschungen zugänglich gemacht haben.<sup>109</sup> Umso mehr habe ich der Firma Orgelbau Waltershausen GmbH in Person von Herrn OBM JOACHIM STADE zu danken. Er ermöglichte mir nicht nur spontan und vollkommen unbürokratisch den Zugang zur ebenso faszinierenden wie berühmten historischen TROST-Orgel in Waltershausen, sondern unterstützte mich auch persönlich nach Kräften mit Rat und Tat bei den Aufnahmen der dortigen *Vox humana* nach Weingartener Vorbild.

Ebenfalls zu danken habe ich Herrn Dr. HANS-WOLFGANG THEOBALD von der Johannes Klais Orgelbau GmbH & Co. KG in Bonn, der neben seiner Aufgabe als Leiter der dortigen Restaurierungsabteilung immer wieder Zeit fand, orgelbauliche und akustische Fragen im Zusammenhang mit der *Vox humana* ausführlich mit mir zu erörtern. Es ist immer wieder faszinierend, auf wie viel Interesse ein doch recht spezielles Thema wie das dieser Dissertation bei unterschiedlichsten Personengruppen stieß und stößt.

Allen genannten und ungenannten Helfern von Herzen Dank!

---

<sup>109</sup>Sehr bedauerlich ist es, dass dies gerade bei einem der wissenschaftlich reizvollsten Instrumente nicht möglich war







# Tabellenverzeichnis

3.1	Artikulationsstellungen für die Vokale A, E, I, O, U nach KRATZENSTEIN. . . . .	39
3.2	Chronologische Erscheinungsreihenfolge der von KEMPELEN referenzierten Grundlagenliteratur zum » <i>Mechanismus</i> «. . .	66
3.3	Bislang bekannte Sprachmaschinen-Nachbauten. . . . .	71
6.1	<i>Windlade</i> mit <i>chromatischer</i> bzw. <i>diatonischer</i> Pfeifen- stellung, Tonfolgen-Schema. . . . .	132
6.2	Tonhöhen der einzelnen VH-Töne (in Hz). . . . .	135
6.3	VH in Amorbach, Meisenheim, Simmern, Waltershausen und Weingarten, Formantwerte für den Ton C. . . . .	136
6.4	Teiltöne mit absolutem Energiemaximum . . . . .	136
6.5	Formantwerte (Hz) $F_1$ und $F_2$ für den VH-Ton $g^0$ in Mei- senheim in verschiedenen Registrierungen. . . . .	142
6.6	Antworten in Prozent für die Stimuli aus Amorbach im Per- zeptionsexperiment Nr. 1. . . . .	150
6.7	Antworten in Prozent für die Stimuli aus Meisenheim im Perzeptionsexperiment Nr. 1. . . . .	150
6.8	Antworten in Prozent für die Stimuli aus Simmern im Per- zeptionsexperiment Nr. 1. . . . .	151
6.9	Antworten in Prozent für die Stimuli aus Waltershausen im Perzeptionsexperiment Nr. 1. . . . .	151
6.10	Antworten in Prozent für die Stimuli aus Simmern im Per- zeptionsexperiment Nr. 2. . . . .	152
6.11	Antworten in Prozent für die Stimuli aus Waltershausen im Perzeptionsexperiment Nr. 2. . . . .	152
6.12	Antworten in Prozent für die originalen VH-Stimuli aus Simmern im Perzeptionsexperiment Nr. 3 (Kategorie I). . .	153

6.13	Antworten in Prozent für die synthetisierten Langvokale nach [Simpson 1998] im Perzeptionsexperiment Nr. 3 (Kategorie II). . . . .	153
6.14	Antworten in Prozent für die resynthetisierten VH-Stimuli aus Simmern im Perzeptionsexperiment Nr. 3 (Kategorie IIIa). . . . .	154
6.15	Antworten in Prozent für die synthetisierten Stimuli mit konstanter $F_0$ und $F_1$ - $F_3$ nach VH (Simmern) im Perzeptionsexperiment Nr. 3 (Kategorie IIIb). . . . .	154
6.16	Antworten in Prozent für die synthetisierten Stimuli mit $F_0$ nach VH (Simmern) und $F_1$ - $F_3$ nach [Simpson 1998] im Perzeptionsexperiment Nr. 3 (Kategorie IIIc). . . . .	155
6.17	Werte für $F_1$ - $F_3$ sowie doppelte und dreifache $F_1$ -Werte für die VH in Simmern und Waltershausen. . . . .	157
8.1	Grundfrequenzen und $F_1$ - $F_3$ (in Hz) der vier <i>Zungenblätter</i> der Saarbrücker SPRACHMASCHINEN-Replik. Die Messungen erfolgten jeweils bei vollständig zusammengesetzter Maschine. . . . .	174

# Abbildungsverzeichnis

2.1	Schematische Darstellung einer klassischen französischen <i>Vox humana</i> . . . . .	11
2.2	GABLERS <i>Vox humana</i> in Weingarten. . . . .	19
2.3	<i>Lingualpfeifen</i> mit <i>aufschlagendem</i> und <i>durchschlagendem Zungenblatt</i> . . . . .	20
3.1	KRATZENSTEINS » <i>Tentamen</i> « (1781), Titelblatt. . . . .	38
3.2	Lage und Positon der Epiglottis beim Menschen in Atem- bzw. Schluckstellung und KRATZENSTEINS Schallgenerator. . . . .	42
3.3	KRATZENSTEINS Vokalresonatoren. . . . .	44
3.4	Formen des Ansatzrohrs zur Vokalartikulation in Kegelschnitten nach KRATZENSTEIN. . . . .	45
3.5	»Sprechende Köpfe« des ABBÉ MICAL. . . . .	55
3.6	KEMPELENS Sprachmaschine bei HINDENBURG (1784) und KEMPELEN (1791) . . . . .	60
3.7	Das »Münchener Exemplar« der KEMPELEN-Sprachmaschine. . . . .	70
4.1	Bestandteile der Versuchsaufbauten von ROBERT WILLIS. . . . .	84
4.2	»Vokalstrahle« zweier unterschiedlicher Konfigurationen der Vokalsynthese von Robert Willis [Willis 1832, Anhang]. . . . .	85
4.3	Die »Artificial voice-box« CHARLES WHEATSTONES. . . . .	87
4.4	Sprachmaschinen-Mundtrichter aus München und London; Kautschuk-Flasche vom Amazonas. . . . .	88
4.5	KEMPELEN-Replik aus München; »Wheatstone's artifical voice box«. . . . .	89
4.6	FABERS »Euphonia«. . . . .	91
4.7	MÄLZELS »sprechende Puppe«. . . . .	99
4.8	KESSELS Stimm-Mechanik für »Papa/Mama«. . . . .	101

5.1	RIESZ' »Mechanical Talker, zwei Schemata. . . . .	104
5.2	»Voder«, Funktionsschema. . . . .	105
5.3	Tibia clausa. . . . .	107
5.4	MILLERS <i>Labialpfeifen</i> -Set für den Vokal /æ/. . . . .	108
5.5	Oszillogramme für /ɑ:/, /ɔ:/, /o:/ und /u:/ von MILLERS Synthese. . . . .	109
5.6	Schematische Schnittzeichnungen der Vokaltrakt-Modelle von CHIBA. . . . .	112
5.7	Vokaltraktmodelle von ARAI nach CHIBA & KAJIYAMA. . .	112
5.8	Vokaltraktmodell von UMEDA & TERANISHI. . . . .	113
5.9	»Waseda Talker No. 7« . . . . .	114
5.10	»Furhat Social Companion Talking Head« . . . . .	117
5.11	Synoptischer Zeitstrahl der mechanischen Sprachsynthese- konzepte von der Antike bis zum Beginn des 20. Jahrhunderts.	121
6.1	Oszillogramme und Spektre von Pfeifen einer historischen und einer modernen <i>Vox humana</i> . . . . .	127
6.2	<i>Vox humana</i> der Orgelbauer STUMM. . . . .	131
6.3	VH in Weingarten, Spektrum des Tons C. . . . .	137
6.4	Amorbach, VH solo, Ton C, Spektrum über 4 Perioden. . .	137
6.5	Meisenheim, VH solo, Ton C, Spektrum über 4 Perioden. .	137
6.6	Simmern, VH solo, Ton C, Spektrum über 4 Perioden. . .	138
6.7	Simmern, VH + Gedeckt 8', Ton C, Spektrum über 4 Peri- oden. . . . .	138
6.8	Waltershausen, VH + Gedeckt 8', Ton C, Spektrum über 4 Perioden. . . . .	138
6.9	Long Term Average Spectra des Tones C der VH aus Amor- bach, Meisenheim, Simmern und Waltershausen. . . . .	139
6.10	Oszillo- und Spektrogramm-Ausschnitte des Tons C der <i>Re-</i> <i>gister</i> VH, <i>Cromorne</i> und <i>Trompette</i> aus Simmern sowie für den Vokal /ø:/. . . . .	140
6.11	Werte für F <sub>1</sub> und F <sub>2</sub> aller aufgenommenen VH-Töne sowie für die Referenzwerte für gesprochene Vokale eines männli- chen Sprechers des Deutschen. . . . .	141
6.12	Werte für F <sub>2</sub> und F <sub>3</sub> aller aufgenommenen VH-Töne sowie für die Referenzwerte für gesprochene Vokale eines männli- chen Sprechers des Deutschen. . . . .	141
7.1	Tabula I aus [Kratzenstein 1781]. . . . .	165
7.2	Tabula II aus [Kratzenstein 1781]. . . . .	166

8.1	Sprachmaschinen-Replik des Autors nach KEMPELEN (2009).	170
8.2	KEMPELENS <i>Zungenpfeife</i> . . . . .	171
8.3	Schematische Zeichnung einer <i>Französischen Löffelkehle</i> (oben) und einer <i>Deutschen Spitzkehle</i> (unten) [nach Bauer 1996, 96] . . . . .	172
8.4	Moderne <i>Lingualpfeife</i> für die Saarbrücker Sprachmaschinen-Replik; vier verschiedene verwendete <i>Zungenblätter</i> . . . . .	174
8.5	<i>Köpfe</i> der Saarbrücker Sprachmaschinen-Replik. . . . .	174
8.6	Spektrale Zusammensetzung der vier <i>Zungenblätter</i> der Saarbrücker SPRACHMASCHINEN-Replik. . . . .	175
8.7	<i>Durchschlagende Harmoniumzunge</i> nach dem <i>Saugwindprinzip</i> . . . . .	178
8.8	<i>Durchschlagende Zungenpfeife</i> mit <i>Harmoniumzunge</i> . . . . .	178
8.9	Zachariaszunge, schematische Schnittzeichnung. . . . .	180
8.10	Schematische Schnittzeichnung und Fotografie eines Doppelrohrblattes. . . . .	181
8.11	Verwendete Doppelrohrblätter: Englischhorn, Dudelsack. . . . .	182
8.12	Membranöse Pfeife. . . . .	184
8.13	Bewertungsraten der Stimuli des zweiten Perzeptionsexperiments. . . . .	188
8.14	Bewertungsraten der Stimuli des dritten Perzeptionsexperiments. . . . .	188
8.15	Schematische Schnittzeichnungen einer <i>Labial-</i> und einer <i>Lingualpfeife</i> . . . . .	LXIV
8.16	Stiftwalze einer mechanischen Orgel. . . . .	LXVII





# Literaturverzeichnis

- [Ablinger 2013] Ablinger, Peter (2013). Speaking Piano (Online-Ressource). *[http://ablinger-mur.at/speaking\\_piano.html](http://ablinger-mur.at/speaking_piano.html)*. (Zuletzt gesehen am 03.09.2013)
- [Adelung 1982] Adelung, Wolfgang (1982). *Einführung in den Orgelbau*. Wiesbaden: Breitkopf.
- [Adlung 1758] \*<sup>110</sup> Adlung, Jakob (1758). *Anleitung zu der musikalischen Gelahrtheit [...]*. Erfurt: Jungnicol.
- Umfangreiches Grundlagenwerk zur Musikgeschichte, Musiktheorie, Musikpraxis sowie zum Orgelbau des 18. Jahrhunderts.*
- [Adlung 1768] Adlung, Jakob (1768). *Musica mechanica organoedi: das ist: Gründlicher Unterricht von der Struktur, Gebrauch und Erhaltung, etc. der Orgeln, Clavicymbel, Clavichordien und anderer Instrumente, in so fern einem Organisten von solchen Sachen etwas zu wissen nöthig ist*. Berlin: Birnstiel.

---

<sup>110</sup>Mit einem Asterisken gekennzeichnete Quellen konnten über den Internet-Service Google Books in digitaler Form eingesehen werden.

*Eine der wichtigsten zeitgenössischen deutschsprachigen Quellen zum Orgelbau der Barockzeit. Basiert in weiten Teilen auf dem zweiten Band von MICHAEL PRAE-TORIUS' »Syntagma musicum« (1619), in dem sich allerdings anders als bei ADELUNG keine Einlassungen zur Vox humana finden.*

[Adolph 2009]

Adolph, Wolfram (2009). »2 Clavier von schwarzem Ebenholtz, die Semitonon aber von Helfenbein...« – Die historische Stumm-Orgel in der evangelischen Schlosskirche St. Wolfgang zu Meisenheim (Glan). *organ – Journal für die Orgel*. Heft 4. S. 12–18.

[Ahrens 2002]

Ahrens, Christian (2002). Zur Frühgeschichte der Instrumente mit Durchschlagzungen in Europa. *Lustig, Monika (Hg.), Michaelsteiner Konferenzberichte*. Bd. 62. Michaelstein. S. 31–50.

[Ahrens 2003]

Ahrens, Christian (2003). Kratzenstein, Gottlob Theophil. *Finscher, Ludwig (Hg.), Die Musik in Geschichte und Gegenwart*. Kassel: Bärenreiter. Personenteil, Bd. 10.

[Ahrens & Braasch 2003]

Ahrens, Christian und Braasch, Jonas (2003). Christian Gottlieb Kratzenstein – De uitvinder van de orgelregisters met doorslagende tongen. *Nederlandse Organistenvereniging (Hg.), Het Orgel*. Heft. 4. S. 32–37.

[Al Moubayed et al. 2013]

Al Moubayed, Samer, Beskow, Jonas und Skantze, Gabriel (2013). The Furhat Social Companion Talking Head. *Proc. 14th Interspeech (Lyon)*. S. 747–749.

[Anonymus 1767]

\*Anonymus (1767). Von einer sogenannten sprechenden Maschine, welche vor einiger

Zeit hier gezeigt wurde. *Hannoverisches Magazin*. Sp. 1649–1652.

[Anonymus 1784a]

Anonymus (1784). Ueber Herrn von Kempelens Schach= Spieler und Sprach= Maschine. Zweeter Brief. *Der Teutsche Merkur*, 1. St., S. 178–182.

[Anonymus 1784b]

Anonymus (1784). Schreiben über die Kempelische Schachspiel- und Redemaschine. *Hessische Beyträge zur Gelehrsamkeit und Kunst*. Bd. 1, Nr. 3, S. 475–487.

[Anonymus 1788]

\*Anonymus (1788). *Ueber Herrn Doctor Müllers redende Maschine und über redende Maschinen überhaupt*. Nürnberg: Gratenhauer.

[Anonymus 1792a]

\*Anonymus (1792). Nachricht von der Sprachmaschine des Herrn Hofr. von Kempelen. *Lichtenberg, Georg Christoph (Hg.), Magazin für das Neueste aus der Physik und Naturgeschichte*. 1792, Bd. 8, 1. Stück. S. 84–102.

*Ausführliche Paraphrase der »V. Abtheilung— Von der Sprachmaschine« aus KEMPELENS »Mechanismus«.*

[Anonymus 1792b]

\*Anonymus (1792). Wien: Wolfgang von Kempelen K. K. wirkl. Hofr. Mechanismus der menschlichen Sprache, nebst der Beschreibung seiner sprechenden Maschine. Wien: Degen 1791. Rezension. *Lichtenberg, Georg Christoph (Hg.), Magazin für das Neueste aus der Physik und Naturgeschichte*. 1792, Bd. 8, 1. Stück, S. 127–142.

[Anonymus 1794]

\*Anonymus (1794). Wunsch. *Reichs-Anzeiger*. Gotha: Becker. Bd. 1. Sp. 703–704

- [Anonymus 1877] Anonymus (1877). Euler, Leonhard. *Historische Commission bei der königl. Akademie der Wissenschaften (Hg.), Allgemeine Deutsche Biographie* Bd 6. Leipzig: Duncker & Humblot. S. 422–431.
- [Anonymus 1883] Anonymus (1883). Kratzenstein, Christian Gottlieb. *Historische Commission bei der königl. Akademie der Wissenschaften (Hg.), Allgemeine Deutsche Biographie*. Bd. 17. Leipzig: Duncker & Humblot. S. 57–61.
- [Anonymus 1900] \*Anonymus (1900). Carl S. Warmholz, ein vergessener Meister und Erfinder. *de Wit, Paul (Hg.), Zeitschrift für Instrumentenbau*. Bd. 20, S. 457–459.
- [Anonymus 1905] \*Anonymus (1905). Die Orgel im neuen Dom zu Berlin. *de Wit, Paul (Hg.), Zeitschrift für Instrumentenbau*. Bd. 25, S. 508–510.
- [Anonymus 2009] Anonymus (2009). Die Stumm-Orgel in Simmern. *Gesellschaft der Orgelfreunde e. V. (Hg.), Ars Organi. Internationale Zeitschrift für das Orgelwesen*. Heft. 2. S. 123.
- [Anonymus 2012a] Anonymus (2012). Dead Media Archive (Online-Ressource). [http://culture-andcommunication.org/deadmedia/-index.php/Euphonia\\_Speaking\\_Machine](http://culture-andcommunication.org/deadmedia/-index.php/Euphonia_Speaking_Machine). (Zuletzt gesichtet am 19.01.2012)
- [Anonymus 2012b] Anonymus (2012). The Theatre Organ Home Page (Online-Ressource). <http://theatreorgans.com/owen-jones/ranks.html>. (Zuletzt gesichtet am 04.09 2012)
- [Anonymus 2013a] Anonymus (2013). Music & Graphic Design (Online-Ressource). <http://mirrormaskcamera.tumblr.com/post/334->

- 86560312/euphonia-speaking-machine-german-inventor-joseph. (Zuletzt gesichtet am 27.06.2013).
- [Anonymus 2013b] Anonymus (2013). Anthropomorphic Talking Robot – Waseda Talker Series (Online-Ressource). <http://www.takanishi.mech.waseda.ac.jp/top/research/voice/>. (Zuletzt gesichtet am 15.06.2013)
- [Anonymus 2013c] Anonymus (2013). Furhat (Online-Ressource). <http://www.speech.kth.se/furhat/sites/default/files/underthehood.jpg>. (Zuletzt gesichtet am 03.09.2013)
- [Arai 2010] Arai, Takayuki (2010). Mechanical Vocal Tract Models for Speech Dynamics. *Proc. Interspeech 2010 (Makuhari)*. S. 1025–1028.
- [Arai Lab 2013] Arai Laboratories (2013). Vocal Tract Models (Online-Ressource). [http://www.splab.net/Vocal\\_Tract\\_Model/index-e.htm](http://www.splab.net/Vocal_Tract_Model/index-e.htm). (Zuletzt gesichtet am 03.09.2013)
- [Ashby & Maidment 2007] Ashby, Michael und Maidment, John (2007). *Introducing Phonetic Science*. Cambridge: University Press.
- [Bärnwick 1922] Bärnwick, Franz (1948). *Die große Orgel im Münster zu Weingarten in Württemberg, erbaut von Joseph Gabler*. Kassel: Bärenreiter.
- [Bauer 1996] Bauer, Siegfried (Hg.) (1996). *Probieren & Studieren. Lehrbuch zur Grundausbildung in der evangelischen Kirchenmusik*. München: Strube.
- [Bedos 1977] Bedos de Celles, Dom François Lamathe / Rensch, Richard (Hg.) (1977). *Die Kunst des Orgelbauers*. Dt. Übers. von Christoph Glatte-Götz. Lauffen: ISO Information.

DOM BEDOS, *Benediktinermönch, Orgelbauer und Mitglied der Pariser Akademie der Wissenschaften, verfasste in deren Auftrag 1766-1778 das detaillierte Grundlagenwerk »L'art de facteur d'orgues« zum klassischen französischen Orgelbau.*

[Biester 1784]

\*Biester, Johann Erich (1784). Schreiben über die Kempelischen Schachspiel- und Redemaschinen. *Biester, Johann Erich (Hg.), Berlinische Monatsschrift*. Bd. 4. S. 495–514.

BIESTER war gemeinsam mit FRIEDRICH NICOLAI und FRIEDRICH GEDIKE die prägende Gestalt der sog. »Berliner Spätaufklärung«. Die von ihm herausgegebene »Berlinische Monatsschrift« war diesen Idealen verpflichtet und gilt aufgrund des in ihr ausgetragenen lebhaften Diskurses um den Begriff der Aufklärung als bevorzugte Zeitschrift IMMANUEL KANTS. Der hier genannte offene Brief befasst sich – wie zeitgleich eine ganze Reihe von Zeitschriftenartikeln – mit einer eingehenden Erläuterung zum Wesen der beiden Maschinen WOLFGANG VON KEMPELENS. BIESTER berichtet u. a. von seiner eigenen Schachtpartie gegen den »Schachtürken«. In einem knappen Nachsatz bewertet er vor dem Hintergrund, dass bislang alles auf diesem Gebiet bekannt gewordene sich über kurz oder lang als Täuschung herausstellt habe, die Sprachmaschine KEMPELENS äußerst kritisch und fordert eine Offenlegung des Mechanismus.

[Bösken 1981]

Bösken, Franz (1981). *Die Orgelbauerfamilie Stumm aus Rhaunen-Sulzbach und ihr*

*Werk: Ein Beitrag zur Geschichte des Orgelbaues am Mittelrhein. Mit Korrekturen und Ergänzungen von Anneliese Bösen.* Mainz: Verlag des Mainzer Altertumsvereins.

[Bois-Reymond 1843]

\*Bois-Reymond, Felix Henri du (1843). *Die Fabersche Sprechmaschine. Beilage zur Königlich privilegierten Berlinischen Zeitung.* 24. Januar 1843.

*Bois-Reymond stellt die zeitgenössische »Euphonia« JOSEPH FABERS in den Kontext der Sprachmaschine WOLFGANG VON KEMPELNS und der als Kempelen-Replik angesprochenen Sprachmaschine von LEONHARD POSCH und bewertet sie als diesen aufgrund ihrer Konstruktion deutlich überlegen.*

[Bois-Reymond 1862]

\*Bois-Reymond, Felix Henri de (1862). *Kadmus, oder Allgemeine Alphabetik.* Berlin: Dümmler.

*In dem die Sprache in den vielfältigen Aspekten ihrer gesprochenen und geschriebenen Form behandelnden Werk wird im Abschnitt über Phonetik den Sprachmaschinen ein eigenes Kapitel gewidmet. Hier referiert der Autor zunächst die Arbeiten WOLFGANG VON KEMPELNS und CHRISTIAN GOTTLIEB KRATZENSTEINS, um folgend auf die zeitgenössischen Sprachmaschinen von LEONHARD POSCH und SALOMON WARMHOLZ einzugehen, mit denen er selbst in Berührung kam.*

[Bonkhoff 1984]

Bonkhoff, Bernhard H. (1984). *Historische Orgeln in der Pfalz.* München: Schnell & Steiner.

- [Braasch 2004] Braasch, Jonas (2004). *Orgelregister mit Durchschlagungen*. Berlin: Logos.
- [Brackhane 2009] Brackhane, Fabian (2009). Perception tests with a replica of von Kempelen's speaking machine. *Proceedings of the NAG/DAGA 2009*. Rotterdam, S. 122-125
- [Brackhane 2011] Brackhane, Fabian (2011). Die Sprachmaschine Wolfgang von Kempelens – Von den Originalen bis zu den Nachbauten. (Gekürzte Fassung der Magisterarbeit). *Trouvain, Jürgen und Barry, William (Hg.), Phonus 16 – Forschungsberichte des Instituts für Phonetik der Universität des Saarlandes*. Saarbrücken. S. 49–148.
- [Brackhane & Trouvain 2008] Brackhane, Fabian und Trouvain, Jürgen (2008). What makes ›mama‹ and ›papa‹ acceptable? Experiments with a replica of von Kempelen's speaking machine. *Proc. 8th International Seminar on Speech Production (ISSP '08)*. Strasbourg, S. 329-332
- [Brackhane & Trouvain 2013a] Brackhane, Fabian und Trouvain, Jürgen (2013). The organ stop "vox humana" as a model for a vowel synthesizer. *Proc. Inter-speech, Lyon*, S. 3172-3176.
- [Brackhane & Trouvain 2013b] Brackhane, Fabian und Trouvain, Jürgen (2013). Stationen Mechanischer Sprachsynthese vom 18. bis zum 20. Jahrhundert. *Mehnert, D., Kordon & U., Wolf, M. (Hg.). Systemtheorie Signalverarbeitung Sprachtechnologie. Rüdiger Hoffmann zum 65. Geburtstag*. Dresden: TUDpress, S. 319-326.
- [Brackhane & Trouvain 2013c] Brackhane, Fabian und Trouvain, Jürgen (2013). On the similarity of tones of the organ stop vox humana to human vowels. *The Phonetician* 107/108, S. 7-20.



[Brewster 1833]

\*Brewster, David (1833). *Briefe über die natürliche Magie an Sir Walter Scott. Übersetzt von Friedrich Wolff*. Berlin: Enslin.

*Der schottische Phyiker, Erfinder sowohl des Kaleidoskops als auch des dioptrischen Stereoskops setzt sich in seinem als Briefsammlung an den schottischen Schriftsteller Sir WALTER SCOTT konzipierten Werk mit vielfältigen Fragestellungen aus Physik und Chemie (seinerzeit noch mit dem Begriff der Magia naturalis bezeichnet) auseinander. Im elften Brief befasst er sich mit mechanischen Automaten seit der Antike und in diesem Zusammenhang auch mit den Arbeiten*

WOLFGANG VON KEMPELENS.  
*Der Übersetzer, der Chemiker FRIEDRICH BENJAMIN WOLFF, war viele Jahre über Professor für Mathematik und Physik am »Joachimsthalschen Gymnasium« in Berlin und Autor eines »Lehrbuchs der Chemie« (1820/21) sowie Mitherausgeber des »Chemischen Wörterbuchs« (1807-1810).*

[Broeke 1983]

Broeke, Marcel van den (1983). Wolfgang von Kempelen's Speaking Machine as a Performer. *Studies for Antonie Cohen*. Foris Publications Dordrecht.

[Bruce 1973]

Bruce, Robert V. (1973). *Bell – Alexander Graham Bell and the conquest of solitude*. Boston: Little.

[Brunner 1798]

\*Brunner, Heinrich Markus (1798). *Ausführliche Beschreibung der Sprachmaschinen oder sprechenden Figuren*. Nürnberg: Zeh.

*Unter dem Kürzel H. M. B. behandelt BRUNNER ausführlich das Phänomen der in seinen Augen immer als Betrug angelegten Sprachmaschinen, das ein deutliches Zeichen dafür sei, wie bereitwillig die Menschheit auf spektakulären Sinnestäuschungen hereinfiele. Er illustriert dies nicht nur mit zahlreichen Berichten über Pseudo-Sprachmaschinen der Vergangenheit, sondern zeigt auch auf, wie er selbst solche Konstruktionen anfertigt und argumentiert, dass er gerade durch diese Tätigkeit besser als jeder andere wisse, dass es eine echte Sprachmaschine niemals geben könne.*

[Burney 1772]

\*Burney, Charles (1772). *Carl Burney's der Musik Doctors Tagebuch einer musikalischen Reise durch Frankreich und Italien*. Hamburg: Bode. Bd. 1.

*Der englische Musikhistoriker und Organist unternahm in den Jahren 1770-1772 für sein Werk »A General History of Music« (1776 ff.) Forschungsreisen auf den europäischen Kontinent. Seine Reiseberichte, in denen er u. a. auch die persönlichen Begegnungen mit CARL PHILIPP EMMANUEL BACH, WOLFGANG AMADEUS MOZART und WILLIBALD GLUCK schildert, erschienen unmittelbar im Anschluss an die Reise sowohl auf Englisch wie auf Deutsch und gelten als einzigartiges Zeugnis des europäischen Musiklebens im späten 18. Jahrhundert.*

[Burney 1773]

\*Burney, Charles (1773). *Carl Burney's der Musik Doctors Tagebuch seiner musikalischen Reisen*. Hamburg: Bode. Bd. 2.

[Busch 1817]

\*Busch, Gabriel Christian Benjamin (1817). Puppe. *Busch, Gabriel Chri-*

*stoph Benjamin (Hg.), Handbuch der Erfindungen.* Eisenach: Bärecke. Bd. 10, I.

*Der evangelische Pfarrer publizierte zwischen 1790 und 1798 ein achtbändiges Kompendium zu bis dahin bekannt gewordenen Erfindungen. Die hier verwendete, ab 1805 erschienene vierte Auflage umfasst bereits zwölf Bände.*

- [Busch 1821] \*Busch, Gabriel Christoph Benjamin (1821). Sprachmaschinen. *Busch, Gabriel Christoph Benjamin (Hg.), Handbuch der Erfindungen* Eisenach: Bärecke. Bd. 11.
- [Busch 2007] \*Busch, Hermann J. und Geuting, Matthias (2007). *Lexikon der Orgel*. Laaber: Laaber.
- [Chapuis & Gélis 1928] Chapuis, Alfred und Gelis, Edouard (1928). *Le Monde des Automates – Etude historique et technique*. Paris: Blondel la Rougery. Bd. 2.
- [Chiba 1941] Chiba, Tsutomu und Kajiyama, Masako (1941). *The vowel. Its nature and structure*. Tokio: Phonetic Soc. of Japan.
- [Darwin 1803] \*Darwin, Erasmus (1803). *The Temple of Nature; or, The Origin of Society*. London: J. Johnson.
- [Darwin 1981] Darwin, Erasmus (1981). *The Letters of Erasmus Darwin*. King-Hele, Desmond (Hg.). Cambridge: Cambridge University Press.
- [Debschitz & Debschitz 2013] Debschitz, Uta von und Debschitz, Thilo von (2013). *Fritz Kahn*. Köln: Taschen.
- [Denker 1889] Denker, Alfred (1889). Bericht über die Versammlung deutscher Ohrenärzte und

- Taubstummenlehrer zu München. *Deutsche Gesellschaft der Hals-Nasen-Ohren-Ärzte (Hg.)*, *Archiv für Ohrenheilkunde*. Bd. 47, S. 198–199.
- [Diderot 1765] Diderot, Denis (Hg.) (1765). *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers*. Band 17. Neufchastel: Faulche.
- [Dodart 1700] \*Dodart, Denis (1700). Sur les causes de la voix de l’homme, et de ses differens tons. *Academie Royale des Sciences (Hg.)*, *Mémoires de l’Academie Royale des Sciences*. Paris.
- [Draxler 2011] Draxler, Christoph (2011). Percy – an HTML5 framework for media rich web experiments on mobile devices. *Proc. 12<sup>th</sup> Interspeech (Florence)*. S. 3339–3340.
- [Dudley & Tarnoczy 1950] Dudley, Homer und Tarnoczy, T. H. (1950). The speaking machine of Wolfgang von Kempelen. *The Journal of the Acoustical Society of America*. Vol. 22. S. 151–166.
- [Eberlein 2007a] Eberlein, Roland (2007). Durchschlagende Zungen. *Busch, Hermann J. und Geuting, Matthias (Hg.)*, *Lexikon der Orgel*. Laaber: Laaber. S. 185–187.
- [Eberlein 2007b] Eberlein, Roland (2007). Vox humana. *Busch, Hermann J. und Geuting, Matthias (Hg.)*, *Lexikon der Orgel*. Laaber: Laaber. S. 817.
- [Eberlein 2009] Eberlein, Roland (2009). *Orgelregister. Ihre Namen und ihre Geschichte*. Köln: Siebenquart.
- [Eberlein 2011] Eberlein, Roland (2011). *Die Geschichte der Orgel*. Köln: Siebenquart.

[Ebert 1785]

Ebert, J. J. (1785). *Nachricht von dem berühmten Schachspieler und der Sprachmaschine des K. K. Hofkammerraths Herrn von Kempelen*. Leipzig: Müllersche Buchhandlung.

*Der Wittenberger Mathematiker und Schriftsteller berichtet in diesem offenen Brief an den Leipziger Rhetorik-Professor AUGUST WILHELM ERNESTI über die beiden aufsehen-erregenden Apparate WOLFGANG VON KEMPELENS nach dessen Auftritt bei der Leipziger Herbstmesse 1784. Wie HINDENBURG, auf dessen Schrift dieser Brief ganz offensichtlich aufbaut, spricht sich EBERT dafür aus, dass gerade die Mangelhaftigkeit der Maschine ein Beweis gegen Betrug sei. Die beigegebene Abbildungstafel der Sprachmaschine ist identisch mit der bei HINDENBURG (1784), nicht jedoch der Text.*

[Edgeworth 1820]

\*Edgeworth, Richard Lovell und Edgeworth, Maria (1820). *Memoirs of Richard Lovell Edgeworth, Esqu. begun by himself and concluded by his daughter, Maria Edgeworth*. London: Hunter. Bd. 2.

[Eidenbenz 2006]

Eidenbenz, Michael (2006). Von Jalousie-, Wind- und anderen Schwellern – die historische Suche nach der Orgel-Expressivität. *Eidenbenz, Michael, Glaus, Daniel & Kraut, Peter (Hg.), Frischer Wind – Die Forschungsorgeln der Hochschule der Künste Bern*. Saarbrücken: Pfau. S. 38–46.

[Ellerhorst 1936]

Ellerhorst, Winfred (1936). *Handbuch der Orgelkunde*. Einsiedeln: Benziger.

*Sehr umfassendes Grundlagenwerk zu technischen Aspekten des Orgelbaues, das jedoch den technologischen und insbesondere ideologischen Stand seiner Zeit widerspiegelt.*

[Eppelsheim 1971]

Eppelsheim, Jürgen (1971). Die Stumm-Orgel der ehemaligen Abtei Amorbach. *Die Musikforschung*. Bd. 24. S. 43-70 und 178-194.

[Essig 2000]

Essig, Rolf-Bernhard (2000). *Der Offene Brief. Geschichte und Funktion einer publizistischen Form von Isokrates bis Günther Grass.*

[Euler 1768]

Euler, Leonhard (1768). *Lettres à une princesse d'Allemagne: sur divers sujets de physique et de philosophie*. Saint Petersburg: Académie Imperiale des Sciences. Bd. 2.

[Euler 1773]

\*Euler, Leonhard (1773). *Briefe an eine deutsche Prinzessinn über verschiedene Gegenstände aus der Physik und Philosophie: Aus dem Französischen übersetzt*. Leipzig: Junius. Bd. 2.

EULER, der einen Gutteil der heutigen mathematischen Symbole prägte, publizierte neben zahlreichen Abhandlungen zu Fragen der Mathematik und Mechanik 1768 die »Lettres à une princesse d'Allemagne«, die bald darauf auch in deutscher Übersetzung erschienen und große Bekanntheit erlangten. Vermittelt werden die Grundzüge von Mathematik, Physik, Astronomie und aber auch Theologie.

[Euler 1862]

\*Euler, Leonhard (1862). *Leonhardi Euleri opera posthuma mathematica et physica*. Petersburg: Eggers. Bd. 2.

*Die erst 1844 aufgefundenen nachgelassenen Schriften EULERS wurden 1862 durch die Petersburger Ademie der Wissenschaften unter der Aegide des Euler-Urenkels und Sekrätters der Akademie, PAUL HEINRICH FUSS veröffentlicht.*

- [Felderer & Strouhal 2004] Felderer, Brigitte und Strouhal, Ernst (2004). *Kempelen – zwei Maschinen*. Wien: Sonderzahl.
- [Fitch & Giedd 1999] Fitch, W. T. und Giedd, J. (1999). *Morphology and developement of the human vocal tract: A study using magnetic resonance imaging*. J. Acc. Soc. Am. 106(3), S. 1511–1522.
- [Flanagan 1965] Flanagan, James L. (1965). *Speech Analysis, Synthesis and Perception*. Berlin: Springer.
- [Förster & Nicolaus 1998] Förster & Nicolaus (1998). *Restaurierung und Rekonstruktion der historischen Stumm-Orgel in der Schlosskirche zu Meissenheim / Glan* (Restaurierungsbericht). Lich: Eigenverlag.
- [Fredriksson 2002a] Fredriksson, Niclas (2002). The History of Free Reeds in Organbuilding, Part 1. *International Society of Organ Builders (Hg.) ISO Journal*. Bd. 14, S. 36–42.
- [Fredriksson 2002b] Fredriksson, Niclas (2002). The History of Free Reeds in Organbuilding, Part 2. *International Society of Organ Builders (Hg.) ISO Journal*. Bd. 15, S. 20–40.
- [Friedrichs 1784] \*Friedrichs, C. F. (1784). Ein Brief aus Paris an einen Freund in Hannover. Ueber den gegenwärtigen Zustand der Meinungen des dasigen Publikums, und die redenden Köpfe des Abbé Mical. *Hannoversches Magazin*. 21. Mai 1784. Sp. 641–653.

*Augenscheinlich plagierende Übersetzung von RIVAROL (1808, 207 ff.), neu datiert und mit F. als Autor bezeichnet. Die im Original ausführlich behandelten Versuche der Brüder MONTGOLFIER werden nur knapp abgehandelt. Spalten 645 ff. über die Sprachmaschinen MICALS und KEMPELENS inhaltlich identisch mit RIVAROL (1808, 229 ff.).*

[Frotscher 1927]

Frotscher, Gotthold (1927). *Die Orgel*. Leipzig: Weber.

[Fukui et al. 2008]

Fukui, Kotaro; Ishikawa, Yuma; Shintaku, Eiji; Ohno, Keisuke; Sakakibara, Nana; Takanishi, Atsuo und Honda, Masaaki (2008). Vocal Cord Model to Control Various Voices for Anthropomorphic Talking Robot. *Proc. 8<sup>th</sup> Int. Seminar Speech Production (Strasbourg)*. S. 341–344.

[Gehler 1825]

\*Gehler, Johann Samuel Traugott (1825). *Automat. Johann Samuel Traugott Gehler's Physikalisches Wörterbuch*. Leipzig: Schwickert. Bd. 1.

*Der Physiker und Jurist veröffentlichte ab 1787 dieses Kompendium des damaligen naturwissenschaftliche Wissens nach kritischen Gesichtspunkten. Nach seinem Tod wurde es zwischen 1825 und 1845 von HEINRICH WILHELM BRANDES et al. in erweiterter Form neu verlegt.*

[Gehler 1836]

\*Gehler, Johann Samuel Traugott (1836). *Schall. Johann Samuel Traugott Gehler's Physikalisches Wörterbuch*. Leipzig: Schwickert. Bd. 8.

[Gehler 1845]

\*Gehler, Johann Samuel Traugott (1845). *Sprachmaschine. Johann Samuel Traugott*



- Gehler's Physikalisches Wörterbuch*. Leipzig: Schwickert. Bd. 11.
- [Gessinger 1994] Gessinger, Joachim (1994). *Auge & Ohr – Studien zur Erforschung der Sprache am Menschen 1700–1850*. Berlin: de Gruyter.
- [Goebel 1975] Goebel, Joseph (1975). *Theorie und Praxis des Orgelpfeifenklanges. Intonieren und Stimmen*. Frankfurt/Main: Verlag das Musikinstrument.
- [Goethe 1893] Goethe, Johann Wolfgang von (1893). *Goethes Werke, herausgegeben im Auftrage der Großherzogin Sophie von Sachsen*. IV. Abtheilung, 12. Band. Weimar: Böhlau.
- [Gorski 2013] Gorski, Markus (2013). Rohrblattinstrumente (Online-Ressource). <http://www.lehrklaenge.de/HTML/rohrblattinstrumente.html> (zuletzt gesichtet am 12.09.2013).
- [Gress 2007] Greß, Frank-Harald (2007). *Die Orgeln Gottfried Silbermanns*. Dresden: Sandstein.
- [Hankins & Silverman 1995] Hankins, Thomas L. und Silverman, Robert J. (1995). *Instruments and the Imagination*. Princeton: Princeton University Press.
- [Helmholtz 1863] Helmholtz, Hermann von (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Braunschweig: Vieweg.
- [Helmont 1657] \*Helmont, Franciscus Mercurius van (1657). *Alphabeti vere naturalis Hebraici brevissima delineatio*. Sulzbach: Lichtenthal.

*Der flämische Mediziner, Diplomat und Kabbalist befasste sich unter anderem auch mit der Frage, ob es eine »Adamitische« oder »Ursprache« gebe. Nach seiner Argumentation handelt es sich hierbei um das Hebräische, da dessen Schriftalphabet wie eine musikalische Partitur die notwendige Artikulationsstellung für den jeweiligen Laut widerspiegeln.*

[Hindenburg 1784]

Hindenburg, Carl Friedrich (1784). *Ueber den Schachspieler des Herrn von Kempelen – nebst einer Beschreibung seiner Sprachmaschine*. Leipzig: Müller.

*Der Leipziger Mathematiker und Physiker beschreibt in seinem offenen Brief an den Schweizer Mathematiker, Physiker und Juristen JOHANN II. BERNOULLI die beiden Maschinen KEMPELENS, so wie sie während der Leipziger Herbstmesse 1784 zu sehen waren. Ebenso wie der im Folgejahr erschienene Bericht JOHANN JACOB EBERTS liegt der Schwerpunkt der Beschreibung auf dem »Schachspieler« und der mathematischen Analyse von dessen Spielweise, während auf die Sprachmaschine erst gegen Ende knapp eingegangen wird, wobei HINDENBURG gerade in der festgestellten Imperfektion der Synthesequalität einen Beweis gegen die Unterstellung eines Betruges sieht.*

[Hoffmann & Mehnert 2007]

Hoffmann, Rüdiger und Mehnert, Dieter (2007). Die Kesselschen Stimm-Mechaniken in der historischen akustisch-phonetischen Sammlung der TU Dresden. *Tagungsband DAGA 2007*. S. 401–402.

[Hoffmann & Löbe 2008]

Hoffmann, Rüdiger und Löbe, Lutz-Peter (2008). Die gemeinsamen Beiträge von

- Johannes Kessel und Ernst Mach zur Entwicklung der Hörakustik. *Tagungsband DAGA 2008*. S. 709–710.
- [Hoffmann 2009] Hoffmann, Rüdiger (2009). »... ein curiöser und inventiöser Mensch« – Johann Valentin Merbitz und sein sprechender Kopf in Dresden. *Hoffmann, Rüdiger (Hg.), Elektronische Sprachsignalverarbeitung 2009. Band 2. Tagungsband des Traditionstages*. Dresden: TUDpress. S. 40–51.
- [Hollingshead 1895] \*Hollingshead, John (1895). *My Lifetime*. Vol. 1. London: S. Low, Marston & C.
- Der englische Theaterimpresario und Schriftsteller beschreibt in seinem Memoiren anschaulich eine Vorführung der »Euphonia« durch JOSEPH FABER.*
- [Holmes 1864] \*Holmes, Oliver Wendell (1864). *Soundings from the Atlantic*. Boston: Ticknor and Fields.
- [Jackson 2005] Jackson, Philip J.B. (2005). Mama and papa: the ancestors of modern-day speech. *Smith, Chr. W. M. (Hg.), The Genius of Erasmus Darwin*. Burlington: Ashgate. S. 217–236.
- [Juškevič & Winter 1959] Juškevič, Adolf P. und Winter, Eduard (1959). *Die Berliner und die Petersburger Akademie der Wissenschaften im Briefwechsel Leonhard Eulers*. Teil 1. Berlin: Akademie-Verlag.
- [Juškevič & Winter 1961] Juškevič, Adolf P. und Winter, Eduard (1961). *Die Berliner und die Petersburger Akademie der Wissenschaften im Briefwechsel Leonhard Eulers*. Teil 2. Berlin: Akademie-Verlag.

- [Juškevič & Winter 1976] Juškevič, Adolf P. und Winter, Eduard (1976). *Die Berliner und die Petersburger Akademie der Wissenschaften im Briefwechsel Leonhard Eulers*. Teil 3. Berlin: Akademie-Verlag.
- [Jones 1935] Jones, Arthur Taber (1935). Organ Pipes and Vowel Quality. *Journal of the American Society of Acoustics*. Bd. 6. S. 282–283.
- [Karsten 1789] \*Karsten, Dietrich Ludwig Gustav (1789). Magnetische Sprechmaschine. *Biester, Johann Erich (Hg.), Berlinische Monatschrift*. Bd. 2. S. 182–188.
- Der Berliner Mineraloge befasst sich in diesem Bericht kritisch mit der scheinbar sprechenden Puppe des DR. MÜLLER. Seine Argumentation gegen eine Echtheit dieser Sprachsynthese fußt neben seiner Skepsis gegenüber den angeblichen Wunderkräften des Magnetismus unter anderem auf der Argumentation, dass sogar der sagenhafte »Schachspieler« KEMPELENS sich als von einem verborgenen Menschen gesteuert erwiesen habe.*
- [Kaufmann 1823] Kaufmann, Friedrich (1823). Ueber die Erfindung der Rohrwerke mit durchschlagenden Zungen. *Allgemeine musikalische Zeitung / Intelligenzblatt*. S. 149–155.
- [Kempelen 1791a] Kempelen, Wolfgang von (1791). *Wolgangs von Kempelen k. k. wirklichen Hofraths Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine*. Wien: Degen.

*Umfassendes Werk zur Sprachforschung, in dem die Sprache sowohl nach philosophischen, (Sprache als Gottesgeschenk oder vom Menschen erworben) als auch physiologischen und artikulatorischen Gesichtspunkten behandelt wird. KEMPELEN stützt sich bei seiner Darstellung neben einer Vielzahl aktueller Publikationen zum Thema auch auf intensive eigene anatomische Beobachtungen und begründet mit diesem Werk faktisch die Disziplin der wissenschaftlich exakten Phonetik. Die abschließende ausführliche Beschreibung seiner Sprachmaschine war bis heute Grundlage für zahlreiche Repliken.*

[Kempelen 1791b]

Kempelen, Wolfgang von (1791). *Le mécanisme de la parole suivi de la description d'une machine parlante*. Wien: Degen.

*Zeitgleich mit der deutschen Ausgabe des »Mechanismus« erschienen und inhaltlich identisch. Der Übersetzer ist unbekannt.*

[Kessel 1900]

Kessel, Johannes (1900). Demonstration von Apparaten zur Erzeugung künstlicher Laute. *Bezold, Friedrich und Passow, Karl Adolf (Hg.), Verhandlungen der Versammlung Deutscher Ohrenärzte und Taubstummenlehrer in München am 16. September 1899*. Berlin: Staude. S. 28–29.

*Der Jenaer Hals-Nasen-Ohren-Arzt, der als Pionier der Mittelohroperationen gilt, stellte eine Serie von »Stimm-Mechaniken« vor, die zur Rehabilitation Hörgeschädigter dienen sollten. Diese basieren auf durchschla-*

*genden Zungenpfeifen in Kombination mit verschiedenen Resonatoren.*

[Kircher 1650]

\*Kircher, Athanasius (1650). *Musurgia universalis*. Tomus II. Romae: Typis Ludouici Grignani.

*Der deutsche Jesuit und Universalgelehrte legt in diesem umfangreichen Werk seine Ansichten zur Musik im Allgemeinen und zur Affektenlehre im Besonderen dar. Ebenso werden Fragen der Schallübertragung und Schallverstärkung behandelt. Zahlreiche Abbildungen.*

[Kircher 1662]

\*Kircher, Athanasius / Andreas Hirsch (Übers.) (1662). *Kircherus Jesuita Germanus Germaniae redonatus: sive Artis Maginae de Consono & Dißono Ars Minor; Das ist/ Philosophischer Extract und Auszug/ aus deß Welt-berühmten Teutschen Jesuitens Athanasii Kircheri von Fulda Musurgia Universali in Sechs Bücher verfasset : Darinnen die gantze Philosophische Lehr und Kunst-Wissenschaft von den Sonis, wie auch der so wol theoretisch- als practischen Music/ mit höchster Varietät geoffenbaret ... und vor Augen gestellt wird / Ausgezogen und verfertiget/ auch mit einem nötigen Indice gezieret von Andrea Hirschen/ Muso- Philo- Sopho- Theo- Philo, Evangel. Pfarrern zu Bächlingen/ in der Grafschaft Hohenloh. Schwäbisch Hall: Laidigen.*

*Stark gekürzte deutsche Übersetzung der »Musurgia universalis«, bei der zudem sämtliche Abbildungen fehlen.*

[Kircher 1684]

Kircher, Athanasius / Nislen, Tobias (Übers.) (1684). *Neue Hall- und Thon-Kunst oder Mechanische Geheim-Verbindung der Kunst und Natur durch Stimme und Hall-Wissenschaft gestiftet*. Ellwangen: Heyl.

*Deutsche Übersetzung der »Phonurgia nova« (1673), in der sich KIRCHER mit zahlreichen Aspekten der Akustik auseinandersetzt. Besondere Aufmerksamkeit erfahren das Echo und Instrumente zur Übertragung und Verstärkung des Schalls. Zusätzlich erörtert KIRCHER der Musik zugeschriebene therapeutische Eigenschaften. Teilweise inhaltliche Wiederholungen aus der »Musurgia universalis«.*

[Kleinert 1989]

Kleinert, Andreas (1989). D'Alembert et le prix de l'académie de Berlin en 1746. *Emery, Monique (Hg.), Jean d'Alembert, savant et philosophe: Portrait à plusieurs voix*. Paris: Éd. des Archives Contemporaines. S. 415–431.

[Kohler 2000]

Kohler, Klaus (2000). The future of phonetics. *Journal of the International Phonetic Association*. Bd. 30, S. 1–24.

[Köhler &amp; Herre 1968]

Köhler, Karl-Heinz und Herre, Gita (Hg.) (1968). *Ludwig van Beethovens Konversationshefte*. Bd. 4.

[Köster 1973]

Köster, Jens-Peter (1973). *Historische Entwicklung von Syntheseparametern zur Erzeugung statischer und vokaltartiger Signale nebst Untersuchungen zur Synthese deutscher Vokale*. Hamburg: Buske.

[Kratzenstein 1781]

Kratzenstein, Christian Gottlieb (1781). *Tentamen resolvendi problema ab Acad.*

*Petropolit. 1780 propositum qualis sit natura litterarum vocalium a, e, i, o, u. Petropoli: Typis Academia Scientiarum.*

*Lateinische Fassung der von der Petersburger Akademie der Wissenschaften prämierten Schrift zum Wesen und zur Synthese der Vokale A, E, I, O, U.*

[Kratzenstein 1782]

Kratzenstein, Christian Gottlieb (1782). Essai sur la naissance & la formation des Voyelles. *Observations sur la physique, sur l'histoire naturelle et sur les arts*. Bd. 21. S. 358–380.

*Französische Übersetzung des »Tentamen« (1781). Übersetzer unbekannt.*

[Krug 1817]

\*Krug von Nidda, Friedrich Albert Franz (1817). Ueber Riestädt und dessen Umgebungen. *Zeitung für die elegante Welt*. 18. August 1817. Sp. 1289–1292.

*Knappe Beschreibung des Flöte spielenden Automaten von CARL SALOMON WARMHOLZ.*

[Kuffner 1808]

\*Kuffner, Christian (1808). Mälzl und seine musikalischen Kunstwerke. *Vaterländische Blätter*. Nr. 14, 24. Juni 1808. S. 112–114.

*Recht pathetische Beschreibung der zahlreichen Musikautomaten JOHANN NEPOMUK MÄLZELS. Keine Erwähnung der von KEMPELEN übernommenen Maschinen.*



[Lampa 1905]

Lampa, Anton (1905). Brief vom 10.08.1905 an die Direction des Museums von Meisterwerken der Naturwissenschaft und Technik, München. *Archiv des Deutschen Museums München*. VA 1806, Physik G-P 1905, sub L.

[Lamy 1699]

\*Lamy, Bernard (1699). *La rhétorique ou l'art de parler*. Amsterdam: Marret.

*Der französische Philosoph und Mathematiker behandelt basierend auf einer grundsätzlichen Erörterung zur Sprachproduktion zahlreiche Aspekte der Rhetorik, wozu auch ein Gedankenexperiment zum Beweis der technischen Unmöglichkeit von Sprachmaschinen zählt.*

[Lee 2013]

Lee, Edward (2013). Parametric Speech Synthesis (Online-Ressource). <http://ptolemy.eecs.berkeley.edu/~eal/audio/voder.html> (zuletzt gesichtet am 16.06.2013).

[Lichtenberg 1783]

\*Lichtenberg, Georg Christoph (1783). o. T. *Lichtenberg, Georg Christoph (Hg.), Magazin für das neueste aus Physik und Naturgeschichte*. Bd. 2, 2. Stück. S. 163–164.

*Notiz über den Antrag des ABBÉ MICAL, seine »Sprechenden Köpfe« von der Pariser Akademié der Wissenschaften prüfen zu lassen mit Auflistung der von der Konstruktion gesprochenen Sätze.*

[Lichtenberg 1784]

\*Lichtenberg, Georg Christoph (1784). o. T. *Lichtenberg, Georg Christoph (Hg.), Magazin für das neueste aus Physik und*

*Naturgeschichte.* Bd. 2, 4. Stück. S. 219–220.

*Notiz über KEMPELENS Sprachmaschine und die Kritik, der diese während der Tournee in England und Frankreich ausgesetzt war.*

[Lichtenberg 1785]

\*Lichtenberg, Georg Christoph (1785). Etwas über den Schachspieler, und die Sprachmaschine des Hrn. Hof-Cammerraths von Kempele. *Lichtenberg, Georg Christoph (Hg.), Magazin für das neueste aus Physik und Naturgeschichte.* Bd. 3, 2. Stück. S. 183–192.

*Ausführliche Überlegungen zur Funktionsweise des »Schachspielers« und Beschreibung der Sprachmaschine und ihrer Bedienung.*

[Liénard 1969]

Liénard, Jean-Sylvain (1969). Reconstruction de la machine parlante de Kempelen. *IV. Budapest Akusztikal Konferencia*, o. S.

[Ljubimenko 1937]

Ljubimenko, I. I. (1937). *Učenaja korespondencija Akademii nauk XVIII veka: Naučnoe opisanie [Gelehrter Briefwechsel der Petersburger Akademie der Wissenschaften im 18. Jahrhundert: Wissenschaftliche Beschreibung]*. Moskva: Izd. Akademii nauk SSSR.

[Locher 1887]

Locher, Carl (1887). *Erklärung der Orgelregister mit Vorschlägen zu wirksamen Register-Mischungen.* Bern: Nydegger & Baumgart.

[Locher 1904]

Locher, Carl (1904). *Die Orgelregister und ihre Klangfarben.* Bern: Nydegger & Baumgart.

- [Locher 1906] Locher, Carl (1906). Über Wert oder Unwert des Zungenregisters "Vox humana". *Urania*. Bd. 63, S. 1–2.
- [Lottermoser 1936] Lottermoser, Werner (1936). *Klanganalytische Untersuchungen an Zungenpfeifen*. Berlin: Junker & Dünnhaupt.
- [Lottermoser 1983a] Lottermoser, Werner (1983) *Orgeln, Kirchen und Akustik*. Band 1. Frankfurt / Main : Bochinsky.
- [Lottermoser 1983b] Lottermoser, Werner (1983). *Orgeln, Kirchen und Akustik*. Band 2. Frankfurt / Main : Bochinsky.
- [Mahrenholz 1928] Mahrenholz, Christhard (1928). Der gegenwärtige Stand der Orgelfrage im Lichte der Orgelgeschichte. *Mahrenholz, Christhard (Hg.), Bericht über die dritte Tagung für deutsche Orgelkunst in Freiberg i. Sa.* 1928.
- [Mahrenholz 1982] Mahrenholz, Christhard (1982). *Die Orgelregister, ihre Geschichte und ihr Bau*. Lauffen: Rensch.
- [Mayer 1788] \*Mayer, Johann Tobias (1788). o. T. *Groß, Johann Heinrich (Hg.), Real-Zeitung auf das Jahr 1788*. Nürnberg / Erlangen.
- Bericht über die scheinbare Sprachmaschine des DR. MÜLLER. Die Beschreibung der Konstruktion unterscheidet sich grundsätzlich von der bei KARSTEN (1789) und ähnelt eher der des »Schachtürken« KEMPELENS.*
- [Meyer 1863] Meyer, Hermann J. (1863). Kautschuk. *Meyer, Hermann J. (Hg.), Neues Konversationslexikon*. Hildburghausen: Bibliographisches Institut. Bd. 9. S. 987–991.

- [Miller 1905] Miller, Oskar von (1905). Brief vom 04.08. 1905 an Prof. Dr. A. Lampa, Wien. *Archiv des Deutschen Museums München*. VA 1806, Physik G-P 1905, sub L.
- [Miller 1922] Miller, Dayton Clarence (1922). *The Science of Musical Sounds*. New York: Macmillan, 1922.
- Der amerikanische Physiker legt in seiner schriftlichen Fassung von acht Vorlesungen aus dem Jahr 1914 dar, wie musikalische Töne aufgezeichnet und analysiert werden können. Das abschließende Kapitel befasst sich mit verschiedenen Möglichkeiten zur Vokalsynthese.*
- [Müller 2008] Müller, Mette (2008). Christian Gottlieb Kratzenstein und die Durchschlagzunge. Pape, Uwe (Hg.), *Georg Joseph Vogler – Umbrüche im Orgelbau*. Bd. 2, S. 141–148.
- [Müller-Heuser 1997] Müller-Heuser, Franz (1997). *Vox humana – Ein Beitrag zur Untersuchung der Stimmmästhetik des Mittelalters*. Kassel: Bosse.
- [Nauwerck 1815] \*Nauwerck, Lebrecht (1815). Die Vervollkommnung des Mechanismus an der deutschen Harfe. *Allgemeine Musicalische Zeitung*. 18. August 1815. Sp. 545–552.
- Erwähnung von CARL SALOMON WARMHOLZ als Entwickler einer neuartigen Harfenmechanik.*
- [Nicolai 1785] \*Nicolai, Friedrich (1785). Die Schachspielende Figur des Hrn. von Kempelen. *Beschreibung einer Reise durch Deutschland und die Schweiz im Jahre 1781. Nebst Bemerkungen über Gelehrsamkeit, Industrie,*

*Religion und Sitten.* Bd. 6. Berlin und Stettin: Eigenverlag. S. 420-436.

*Der Berliner Schriftsteller und Hauptvertreter der »Berliner Aufklärung« legt seine kritischen Ansichten über den »Schachspieler« WOLFGANG VON KEMPELENS dar, den er selbst vergeblich zu besichtigen versucht hatte. Ähnlich ablehnend äußert er sich über KEMPELENS Sprachmaschine, ohne von ihr mehr zu kennen als die Beschreibung bei [Hindenburg 1784]. Ihre Funktionsweise missversteht er hierbei grundlegend.*

- [Niemann 1920] Niemann, W. (1920). Sprechende Figuren – Ein Beitrag zur Vorgeschichte des Phonographen. *Geschichtsblätter für Technik und Industrie.* Bd. VII, S. 2–31.
- [Nikléczy & Olaszy 2003] Nikléczy, Péter und Olaszy, Gabor (2003). Kempelen's speaking machine from 1791: Possibilities and limitations (recovering a 200 year-old technology). *Grazer Linguistische Studien* 62, S. 111-120.
- [Ohala 2011] Ohala, John (2011). Christian Gottlieb Kratzenstein: Pioneer in speech synthesis. *Proc. 17th ICPhS (Hong Kong)*, S. 156-159.
- [Paget 1930] Paget, Richard (1930). *Human Speech.* London: Toutledge and Kegan Paul.
- [Pétursson & Neppert 2002] Pétursson, Magnús und Neppert, Joachim M. H. (2002). *Elementarbuch der Phonetik.* Hamburg: Buske.
- [Poggendorff 1843] \*Poggendorff, Johann Christian (1843). XXII. Faber'sche Sprechmaschine. *Poggendorff, Johann Christian (Hg.), Annalen der Physik und Chemie.* Zweite Reihe, Bd. 28, Heft 1. S. 175–176.

*Notiz über die »Euphonia«* JOSEPH FABERS,  
*Vergleich mit der Sprachmaschine* WOLFGANG VON KEMPELENS.

- [Pompino-Marschall 1991] Pompino-Marschall, Bernd (1991). Wolfgang von Kempelen und seine Sprechmaschine – Eine biographische Notiz zum 200. Jahrestag der Publikation seines »Mechanismus der menschlichen Sprache«. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München*. Bd. 29, S. 181–252.
- [Pompino-Marschall 2004] Pompino-Marschall, Bernd (2004). Von Kempelen's contribution to the theory of acoustic articulation. *Grazer Linguistische Studien*. Bd. 62, S. 137–147.
- [Pompino-Marschall 2005] Pompino-Marschall, Bernd (2005). Von Kempelen et al. – Remarks on the history of articulatory-acoustic modelling. *ZAS Papers in Linguistics*. Bd. 20, S. 145–159.
- [Pompino-Marschall 2009] Pompino-Marschall, Bernd (2009). *Einführung in die Phonetik* (3. Auflage). Berlin: de Gruyter.
- [Reininger 2007] Reininger, Alice (2007). *Wolfgang von Kempelen – Eine Biografie*. Wien: Praesens.
- [Reyher 1693] \*Reyher, Samuel (1693). *De natura et jure auditus ac soni*. Kiel: Reumann.
- Der Mathematiker und Astronom erläutert seine Ansichten über den Schall und das menschliche Gehör.*
- [Reko 1911] Reko, Victor A. (1911). Neue Dokumente zur Geschichte der Sprechmaschine. *Phonographische Zeitschrift*. 12. Jg., No. 18, S. 415–425.

[Rivarol 1808]

\*Rivarol, Antoine de (1808). *Oeuvres complètes de Rivarol*. Paris: Collin. Bd. 2.

*Hierin offener Brief vom 20. 09. 1783 an einen Anonymus, in dem er sich hauptsächlich mit den Versuchen der Gebrüder Montgolfier befasst. Auf den Seiten 229-236 werden die zu diesem Zeitpunkt gerade der Pariser Akademie der Wissenschaften vorgestellten »Têtes parlantes« des ABBÉ MICAL eingehend geschildert. In einer umfangreichen Fußnote S. 242-246 wendet sich RIVAROL polemisch gegen WOLFGANG VON KEMPELEN, der sich zeitgleich mit seinem beiden Maschinen in Paris aufhielt.*

[Rm 1784]

Rm (1784). Windisch, K. G. v.: Briefe über den Schachspieler des Hrn. von Kempele: Rezension. *Allgemeine deutsche Bibliothek*. Bd. 58 (1), S. 275-280.

*Hinter dem Pseudonym verbirgt sich gemäß dem Karlsruher virtuellen Katalog (KVK) vermutlich H. F. VON DIETZ. Die von WINDISCH ebenso wie beim »Schachspieler« einigermmaßen euphorisch gehaltene Beschreibung der Sprachmaschine KEMPELENS findet in dieser Rezension einen polemischen Gegenpart. Wie zeitgleich NICOLAI missversteht der Autor die Funktionsweise der Sprachmaschine grundlegend und begreift sie als angeblich autonom wahrnehmenden und handelnden Automaten.*

[Rohlf 1999]

Rohlf, Johannes (1999). Die Zachariaszunge. Das erste Orgelregister mit einer »gewendeten Zunge«. *ISO Journal* Nr. 6. S. 40-42.

- [Rotermund 1810] \*Rotermund, Heinrich Wilhelm (1810). Von Kempelen. *Fortsetzung und Ergänzungen zu Christian Gottlieb Jöchers allgemeinem Gelehrten-Lexikon*. Leipzig: Gleditsch. Bd. 3. Sp. 211–212.
- [Savart 1826] \*Savart, Félix (1826). Memoir on the Mechanism of the Human Voice. *Brewster, David (Hg.), The Edinburgh Journal of Science*. London: Lumley, 1826, Bd. 4. S. 80–85.
- [Savoy 2010] Savoy, Bénédicte (2010). *Kunstraub – Napoleons Konfiszierungen in Deutschland und die europäischen Folgen*. Berlin: Böhlaus.
- [Schnorr von Carolsfeld 1885] \*Schnorr von Carolsfeld, Franz (1885). Merbitz, Johann Valentin. *Historische Commission bei der königl. Akademie der Wissenschaften (Hg.), Allgemeine Deutsche Biographie*. B. 21. Leipzig: Duncker & Humblot. S. 384–385.
- [Schumann 1940] Schumann, Paul (1940). *Geschichte des Taubstummenwesens*. Frankfurt (Main): Diesterweg.
- [Simpson 1998] Simpson, Adrian (1998). *Phonetische Datenbanken des Deutschen in der empirischen Sprachforschung und der phonetischen Theoriebildung*. Kohler, Klaus (Hg.), Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK) 33. Kiel : Inst. für Phonetik und Digitale Sprachverarbeitung.
- [Sobotta 1959] Sobotta, Johannes (1959). *Atlas der deskriptiven Anatomie des Menschen*. Bd. 2. Berlin: Urban & Schwarzenberg.
- [Splinter 2007] Splinter, Susan (2005). *Zwischen Nützlichkeit und Nachahmung – Eine Biographie*



- des Gelehrten Christian Gottlieb Kratzenstein (1723–1795)*. Frankfurt (Main): Lang.
- [Stevens 1998] Stevens, Kenneth N. (1998). *Acoustic Phonetics*. Cambridge & London: MIT Press.
- [Strong 2007] Strong, Roy (2007). *A little history of the english country church*. London: Vintage Books.
- [Techmer 1884] Techmer, Friedrich H. H. (1884). Naturwissenschaftliche Analyse und Synthese der hörbaren Sprache. *Internationale Zeitschrift für allgemeine Sprachwissenschaft*. S. 67–195.
- [Theobald 1983] Theobald, Hans-Wolfgang (1983). Die Orgel der ehemaligen Abteikirche Amorbach. *Gesellschaft der Orgelfreunde e. V. (Hg.), Ars Organi. Internationale Zeitschrift für das Orgelwesen*. Heft 3, S. 186–192.
- [Thornton 1980] Thornton, Barbara (1980). Vokale und Gesangstechnik: Das Stimmideal der aquitanischen Polyphonie. *Reidemeister, Peter (Hg.), Basler Jahrbuch für historische Musikpraxis*. 4, S. 133–150.
- [Trouvain & Brackhane 2010] Trouvain, Jürgen und Brackhane, Fabian (2009). Zur heutigen Bedeutung der Sprechmaschine Wolfgang von Kempelens. *Elektronische Sprachsignalverarbeitung 2009*, Bd. 2 (Tagungsband des Traditionstages Dresden, 23./24. September 2009), S. 97–107
- [Trouvain & Brackhane 2011a] Trouvain, Jürgen und Brackhane, Fabian (2011). The Relevance Today of Wolfgang von Kempelens's ›Speaking Machine‹. *Trouvain, Jürgen und Barry, William (Hg.), Phonus 16 – Forschungsberichte des Instituts für Phonetik der Universität des Saarlandes*. Saarbrücken, S. 149–166

- [Trouvain & Brackhane 2011b] Trouvain, Jürgen. und Brackhane, Fabian (2011). Wolfgang von Kempelen's speaking machine as an instrument for demonstration and research. *Proc. 17th International Congress of Phonetic Sciences*. Hong Kong. S. 164-167.
- [Vogler 1810] Vogler, Georg Joseph (1810). Ueber Sprach- und Gesang-Automaten. *Sammlung einiger in dem Frankfurter Museum vorgetragenen Arbeiten*. Heft 1. S. 118-130. Frankfurt.
- Der Musiktheoretiker und Organist VOGLER vergleicht Kempelens Sprachmaschine mit dem Orgelregister »Vox humana« und unterscheidet hierbei zwischen reiner »Diction« und echter »Declamation«.*
- [Vogler o. J.] Vogler, Georg Joseph (o. J.). *Abhandlung über die Simplifikation des Orgelbaues* (unpaginiertes Manuskript).
- [Wan et al. 2013] Wan, Robert et al. (2013). Photo-realistic expressive text to talking head synthesis. *Proc. 14th Interspeech (Lyon)*, S. 2667–2669.
- [Weber & Smits 2003] Weber, A. und Smits, R. (2003). Consonant and vowel confusion patterns by American English listeners. *Proc. 15th. ICPHS (Barcelona)*, S. 1437–1440.
- [Wheatstone 1879] Wheatstone, Charles (1879). Reed organ-pipes, speaking machines, etc. *The Physical Society of London (Hg.), The Scientific Papers of Sir Charles Wheatstone, D.C.L., F.R.S.* London: Taylor and Francis, 1879, S. 348–367.
- [Wilke 1825] Wilke, Friedrich (1825). Nachtrag über die Erfindung der Rohrwerke mit durchschla-

- genden (freyschwingenden) Zungen. *Allgemeine musikalische Zeitschrift / Intelligenzblatt*. Bd. 27, S. 263–266.
- [Willis 1832] Willis, Robert (1832). Ueber Vocaltöne und Zungenpfeifen. *Poggendorff, Johann Christian (Hg.), Annalen der Physik und Chemie*. Drittes Stück, S. 397–437.
- [Windisch 1783a] \*Windisch, Karl Gottlieb (1783). *Karl Gottlieb von Windisch's Briefe über den Schachspieler des Herrn von Kempelen nebst drey Kupferstichen die diese berühmte Maschine vorstellen, herausgegeben von Chr. von Mechel der K. K. und anderer Akademien Mitglieder*. Basel: Mechel.
- [Windisch 1783b] \*Windisch, Karl Gottlieb (1783). *Briefe über den Schachspieler des Herrn von Kempelen*. Preßburg: Löwe.
- [Young 1845] \*Young, Thomas (1845). *A course of lectures on natural philosophy and the mechanical arts*. Volume 2. London: Taylor and Walton.

*Der Augenarzt und Physiker YOUNG stellte bereits 1807 eine Reihe von ausformulierten Vorlesungen über zahlreiche Aspekte der Physik zusammen, die posthum erneut veröffentlicht wurden.*

*Im Kapitel XXXIV über die Akustik von Musikinstrumenten behandelt er knapp sowohl die menschliche Stimme und das Orgelregister »Vox humana« wie auch die Sprachsynthese-Versuche KRATZENSTEINS und KEMPELENS.*

# Index

- Ablinger, Peter, 6, 116  
 Adelung, Wolfgang, 126  
 Adlung, Jakob, 10, 11, 15  
 Ahrens, Christian, 46, 47  
 Akademie der Wissenschaften, fran-  
     zösische, 11, 54, 55, 57, 60  
 Akademie der Wissenschaften, könig-  
     lich-preußische, 37  
 Akademie der Wissenschaften, rus-  
     sische, 26, 32, 35–38, 48, 50  
 Alembert, Jean-Baptiste le Rond d',  
     37  
 Amorbach (Odenw.), Benediktiner-  
     Abtei, 130, 131, 134, 135,  
     156  
 Anthon, Hr., 67  
 Anthropoglossa (Register), 12, 14  
 Arai, Takayuki, 111  
 Azyr, Félix Vicq d', 54, 66  
  
 Bach, Wilhelm Friedemann, XIII, XV,  
     15, 16  
 Bacon, Robert, 119  
 Bacon, Roger, 3, 57, 111  
 Bauer, Dominik, 169  
 Bedos de Celles, Dom François L.,  
     11  
 Beethoven, Ludwig van, 98  
 Belcanto, 16, 17  
 Bell Laboratories, 103, 106  
 Bell, Alexander Graham, 26, 89, 96,  
     111  
 Bell, Alexander Melville, 96  
 Bell, Melville, 96  
 Bernoulli, Daniel, 37  
 Bernoulli, Johann II., 32  
 Bernoulli-Prinzip, 20  
 Bois-Reymond, Félix-Henri de, 79,  
     82, 91, 94  
 Brewster, David, XIV, 50, 64, 65  
 Brosset, Charles de, 66  
 Brunner, Heinrich Markus, 78  
 Burney, Charles, 15, 17  
 Busch, Gabriel Christoph Benjamin,  
     3, 31  
  
 Canto a tenore, 16  
 Chiba, Tsutomu, 111  
 Court de Gébelin, Antoine, 66  
 Cromorne (Register), 132  
  
 Darwin, Erasmus, 29, 31, 37, 78  
 Denon, Dominique Vivant, 81, 82  
 Descartes, René, 59  
 Deutsches Museum (München), 69,  
     87, 88  
 Diderot, Denis, 11, 59  
 Dodart, Denis, 23, 57, 66, 183  
 Doppelrohrblatt, 170, 180  
 Dudley, Homer, 106  
  
 Eberlein, Roland, 13  
 Eisleben, 79  
 Ellerhorst, Winfried, 142  
 Epée, Charles-Michel de l', 59

- Euler, Johann Albrecht, 36, 37, 50  
 Euler, Leonhard, 6, 22, 24, 31–33, 35–37, 57, 67, 96, 120  
 Eulersche Zahl, 32  
  
 Faber, Joseph, 26, 90, 93, 94, 96, 193  
 Ferkel, Anton, 67  
 Ferrein, Antoine, 66  
 Franklin, Sir Benjamin, 29, 54  
 Fukui, Kataro, 113  
  
 Gabler, Joseph (Orgelbauer), 17–19, 143  
 Gedeckt (Register), 10  
 Gleib, Reinhold, LVII  
 Goethe, Johann Wolfgang von, 68  
 Golem, 3  
 Gottsched, Johann Christoph, 59  
 Grenié, Gabriel Joseph (Orgelbauer), 40  
 Gress, Frank-Harald, 14  
 Gummiflasche (Kautschuk-), 87  
  
 Haller, Johann Albrecht von, 40, 66  
 Harmonium, 178  
 Hawking, Lucy, XIII  
 Hawking, Stephen, XIII  
 Heinz Nixdorf MuseumsForum (Paderborn), 169  
 Helmholtz, Hermann von, 57, 183, 194  
 Helmont, Franciscus Mercurius van, 63, 66  
 Henry, Joseph, 95  
 Herder, Johann Gottfried von, 66  
 Hindenburg, Carl Friedrich, 59  
 Hjärne, Urban, 5  
 Hoffmann, Rüdiger, 4, 100  
 Hollingshead, John, 94  
 Hugo Mayer Orgelbau GmbH, 169  
 Humboldt, Wilhelm von, 82  
  
 Institut für Deutsche Sprache (Mannheim), 144  
  
 Jäger, Gustav, 69  
 Jones, Arthur Taber, 110  
  
 K. K. technische Hochschule (Wien), 69, 70  
 Kalmar, Jiří, 66  
 Kahn, Fritz, 35  
 Kajiyama, Masato, 111  
 Kautschuk, 64, 87, 90  
 Kehlgesang, 16  
 Kempelen, Wolfgang von, XIV–XVI, 4, 21–23, 26, 27, 29, 31, 33, 37, 42, 44, 46, 48, 54, 57–61, 64–68, 75, 78–83, 85–87, 89, 90, 93, 96–98, 103, 108, 109, 117, 119, 164, 169–172, 176, 177, 180, 181, 183, 184, 193–195  
 Kessel, Johannes, 100  
 King's College (London), 86  
 Kircher, Athanasius, 4, 119, 193  
 Kirschnick (Kirschnick), Franz, 47, 48, 63  
 Köster, Jens Peter, 55  
 Konservatorium (Wien), 70  
 Kopenhagen, 32, 48  
 Kopf, sprechender, 3–5, 52, 53, 55, 57, 97, 113  
 Kratzenstein, Christian Gottlieb, XIV, XV, 21–23, 25, 26, 29, 31, 32, 36–38, 41–44, 46–51, 56, 57, 59, 66, 84, 86, 120, 163, 171, 173, 176, 177, 183  
 Krebs, Johann Ludwig, 15, 16  
 Kunstammer, königlich-preußische, 82, 91  
  
 Labialpfeife, 106, 110, 156, 196

- Lalande, Joseph Jérôme Lefrançois de, 50  
 Lampa, Anton, 69  
 Lamy, Bernard, 6, 34  
 Laplace, Pierre-Simon, 54  
 Lavoisier, Antoine Laurent de, 54  
 Leopoldina (Deutsche Akademie der Naturforscher), 32  
 Lichtenberg, Georg Christoph, 54, 61  
 Lieutaud, Joseph, 66  
 Ligeti, György, 35  
 Lingualpfeife (Zungen-), 9, 13, 19–22, 41, 46, 84, 126, 129, 132, 173, 176, 181, 194  
 Lingualpfeife, aufschlagende, 21, 22, 56, 66, 106, 170, 171, 177, 178, 180, 193  
 Lingualpfeife, durchschlagende, 21, 22, 46–48, 85, 164, 176–181  
 Linné, Carl von, 32  
 Liscovius, Karl Friedrich Salomon, 41  
 Locher, Carl, 13  
 London, 59, 96  
 Lottermoser, Werner, 18, 125, 127, 130, 135, 136, 156, 195  
 Lunar Society (Birmingham), 30  
 Mälzel, Johann Nepomuk, 67, 98  
 Müller, L., 75, 76  
 Müller-Heuser, Franz, 17  
 Magnus, Albertus, 3, 111, 119  
 Mahrenholz, Christhard, 14  
 Maria Theresia von Österreich, 58  
 Meckel, Philipp Th. (?), 40  
 Meisenheim (Glan), ev. Schlosskirche, 130, 134, 135, 143  
 Merbitz, Johann Valentin, 4  
 Mettrie, Julien Offray de la, 59  
 Mical, Abbé, XIX, 26, 29, 31, 37, 52–54, 57, 60, 78, 117  
 Mikoletzky, Juliane, 70  
 Miller, Dayton Clarence, 106, 107, 109  
 Miller, Oskar von, 69, 88  
 Moritz, Karl Philipp, 59  
 Moscheles, Ignaz, 98  
 Mundorgel, 22, 46  
 Nachahmungstheorie, 14  
 Neues Museum (Berlin), 82  
 Newton, Sir Isaac, 44  
 Nicolai, Friedrich, 63  
 Niemann, W., 26  
 Obertonsingen, 16  
 Ohala, John, 85, 164  
 Orchestrion, 48  
 Orgel, 5, 6, 24, 33, 34, 36, 84, 104  
 Paget, Sir Richard, 86  
 Paris, 6, 50, 51, 54, 57, 59–61, 81, 82, 94  
 Pasch, Georg, 6  
 Perzeptionstest, 143  
 Pfeife, membranöse, 57, 66, 170  
 Pfeifenstellung, chromatische, 131, 132  
 Pfeifenstellung, diatonische, 131, 132  
 Posch, Leonhard, 81, 91  
 Praetorius, Michael, 17  
 Rackwitz, Christoffer, 48, 51  
 Ravalement, 53, 54  
 Regal (Register), XIV, 9, 13, 14, 21–23  
 Reiniger, Alice, 58  
 Resonator, LX, LXII, 10, 12, 14, 18, 20–23, 34, 35, 44–46, 50, 55, 56, 84, 85, 126, 128, 130, 132–134, 142, 156, 157, 164, 170, 179, 180  
 Reyher, Samuel, 5  
 Riesz, Robert R., 103, 193

- Ritsch, Winfried, 116  
 Rivarol, Antoine de, 52, 54, 56, 60  
 Robert, Etienne-Gaspard (Etienne Robertson), 51  
 Roy, Charles-George le, 54  
 Russel, G. Oscar, 110  
  
 Süßmilch, Johann Peter, 34, 66  
 Sardinien, 17  
 Savart, Félix, 183  
 Schachspieler, mechanischer (Schachtürke), 58, 59, 67, 75  
 Schneider, Joachim, 16  
 Schubart, Christian F. D., 15, 130  
 Schumacher, Johann Daniel, 37  
 Science Museum (London), 86  
 Siemens, Werner von, 86  
 Silbermann, Gottfried (Orgelbauer), 15  
 Simmern (Hunsrück), ev. Stadtkirche, 130, 134, 135, 139, 144, 148, 156, 158, 159  
 Smits, R., 158  
 Sprachmaschine, XIX, 6, 7, 33, 78  
 Sprachmaschine (Bell), 96, 97  
 Sprachmaschine (Darwin), 29  
 Sprachmaschine (Deutsches Museum, München), 87–89  
 Sprachmaschine (Faber), 26, 79, 89, 95, 96  
 Sprachmaschine (Kempelen), 21, 23, 57–59, 61, 64, 65, 67, 68, 75, 77–79, 81, 82, 93, 96, 98, 104, 109, 169, 176, 177, 184, 194, 195  
 Sprachmaschine (Müller), 75  
 Sprachmaschine (Posch), 81, 82  
 Sprachmaschine (Riesz), 103  
 Sprachmaschine (Warmholz), 79, 80  
 Sprachmaschine (Wheatstone), 86–88, 95, 96  
 Sprechmaschine, XIX  
 Sprechmaschine (Mical), 57  
 St. Petersburg, 32, 46, 48, 50, 59, 66  
 Staehlin, Jacob von (Jakow Jakowlewitsch St.), 35  
 Stiftwalze, 54  
 Stimmideal (Gesang), 16, 17  
 Stimmtonhöhe, 134  
 Stork (Storch), Johann Friedrich, 59  
 Stumm, Gebr. (Orgelbauer), 10, 15, 130, 131, 135, 142, 143, 156, 159  
  
 Technische Universität (Wien), 69, 70  
 Teranishi, R., 111  
 Tremulant, 10, 126  
 Trompette (Register), 132  
 Trost, Tobias H. G. (Orgelbauer), 130  
  
 Umeda, Noriko, 111  
 Universität Wien, 69  
  
 Vaucanson, Jacques de, 52  
 Voce umana (Register), 14  
 Vocoder, 106  
 Voder, 106  
 Vogler, Abbé Georg Joseph, 47, 48, 98  
 Vokalorgel (Kratzenstein), 36, 43, 49, 50, 163, 164, 173, 177  
 Vox humana (Register), XV, LX, 5, 9–11, 13–18, 21–27, 34, 37, 41, 42, 46, 66, 84, 125, 126, 128–131, 135, 139, 142, 143, 156, 158–160, 170, 173, 195  
 Vox humana (Weingarten), 19  
 Vox humana, Sage von der, 17

Waltershausen (Thür.), ev. Stadtkirche, 130, 131, 135, 142–145, 159  
Warmholz, Carl Salomon, 79, 80  
Waseda Talker, 113–115  
Weingarten, Benediktiner-Abtei, 18, 19, 130, 135  
Wernigerode, 31  
Wheatstone's artificial voice box, 26  
Wheatstone, Sir Charles, 5, 26, 27, 69, 86, 89, 90, 95, 96  
Wilkins, John, 66  
Willis, Robert, 23, 27, 37, 45, 48, 83, 85, 86, 177  
Windisch, Karl Gottlieb, 58  
Wolff, Friedrich Benjamin, 65  
  
Zacharias, Ernst, 179  
Zachariaszunge, 179  
Zobel, Rudolf Heinrich, 66  
Zooglossa (Register), 12, 14



# Appendix

## Auszüge aus Kratzensteins »*Tentamen*«

Im Folgenden werden die vier Paragraphen des »*Tentamen*«, die für das Verständnis der Konstruktion von KRATZENSTEINS »Vokalorgel« besonders wichtig sind, ungekürzt wiedergegeben. Die Übertragung ins Deutsche wurde hierbei auf vergleichender Grundlage der lateinischen und französischen Originalfassungen (Kratzenstein 1781; Kratzenstein 1782) besorgt.<sup>111</sup> Der Übersetzung der §§ 15, 24 und 25 liegt eine Fassung von Prof. Dr. REINHOLD GLEI, Bochum, aus dem Lateinischen zugrunde, die vom Autor überarbeitet und kommentiert wurde.<sup>112</sup> Notwendige Erläuterungen und Präzisierungen wurden in eckigen Klammern ergänzt.

### § 15.

»Man stelle eine Röhre AB gemäß Fig. 1 aus Metall oder Elfenbein her, von beliebiger Länge mit einem Durchmesser von 9<sup>'''</sup> [Linien, 1<sup>'''</sup>=2,26 mm], die mit einem gewölbten Endstück BCD wie mit einem Kehlkopf ausgestattet ist, deren Mündung CD zu C hin wie bei einem aufgesetzten Mundstück ein wenig verengt ist, während der Rand an dem gegenüberliegenden Seite eben ausläuft, so dass der Schnitt der Öffnung CD wie eine Parabel EhF ist (gemäß Fig. 2), deren Achse kh der Linie gemäß Fig 1 entspricht. Diese Öffnung soll mit einem Plättchen [dem Zungenblatt] aus Metall, Elfenbein oder Horn – wie es beliebt – versehen werden, wobei es aber mit dem Hals an der ebenen Seite D so verbunden werden muss, dass es an der Verbindungsstelle elastisch wippen kann – einerseits durch das

---

<sup>111</sup>Die inhaltliche Übereinstimmung der Quellen wurde hierbei überprüft.

<sup>112</sup>Mein Dank gilt Prof. GLEI für die Überlassung der unveröffentlichten Übersetzung.

Material des Plättchens selbst, andererseits durch äußere Einwirkung. Außerdem muss dieses Plättchen genau in die Mündung [CD] passen, und zwar so, dass es, wenn es gebogen wird, sich ohne Reibung auf und ab bewegen kann. Diesem Plättchen soll im Ruhezustand von der Seite aus, an der es befestigt ist, die Lage kh zukommen, so dass es ungefähr bis zu einer halben Linie [also ca. 1,1 mm] unterhalb des Randes D abgesenkt ist. Wenn man nun in die Röhre AB hineinbläst, wird die Luft das Plättchen in schwingende Bewegung entlang des kleinen Bogens ki versetzen und – auf diese Weise in Schallwellen geteilt – einen Klang und einen Ton erzeugen, der der menschlichen Stimme ähnlich ist. Die Tonhöhe wird vom Grad der Elastizität der Verbindung in k und von der Masse des Plättchens anhängen; nach diesem Grad [der Elastizität] muss auch die Schwingungsweite Ch bemessen werden. Ein solches Instrument wird Bestandteil der ›Vokalerzeugungsmaschine‹ sein, die an die Akademie übersandt werden wird, damit die tatsächliche Ähnlichkeit [der Töne mit der menschlichen Stimme] im Experiment nachvollzogen werden kann.«

## § 24.

»Da die Herstellung dieser [im vorangehenden Paragraph beschriebenen] Kehle [gemeint ist eine herkömmliche Zungenpfeife] den Orgelbauern bekannt ist und ihnen ein wenig leichter fällt als die Herstellung der [in § 15 beschriebenen] neuen, von mir beschriebenen Kehle, will ich mit wenigen Worten darlegen, wie die meisten Fehler der alten Kehle [der bisherigen Konstruktion der *Zungenpfeifen* mit *aufschlagende Zungenblättern*] vermieden werden können; dadurch wird jene neue [erneuerte *Kehle*] zur hinreichend deutlichen Hervorbringung einiger Vokale ausreichen. [Diese Maßnahmen sind] nämlich: 1) Jener Teil der Kehle bzw. der halbzylindrischen Röhre, der in einem Holzfuß [dem *Kopf*] befestigt ist, muss mit einer bronzenen Querplatte bedeckt und mit flüssigem Zinn fest damit verlötet werden, und zwar so, dass dessen oberster Teil präzise auf derselben Ebene mit dem Rand des Mundstücks [der *Kehle*] liegt; so kann nämlich der hölzerne Keil, der auf das Plättchen gesetzt wird, dieses nicht nach unten biegen [vgl. Abb. 7.2, Fig. 5 und 6]. 2) Der Rand des Mundstücks, der über einem Wetzstein plan geschliffen wurde, muss gut geglättet und dann

über einem geraden Balken aus Hartholz mit Hilfe von tripolitatischer Erde [Quarzsand], die mit viel Öl versetzt wurde, gründlich poliert werden. Die Politur verhindert nämlich über viele Jahre, dass Feuchtigkeit am Rand haften und dass Grünspan ansetzen kann. 3) Dasselbe muss hinsichtlich der unteren Seite des Plättchens [des *Zungenblattes*] gemacht werden. 4) Die Krümmung des Plättchens muss so gestaltet sein, dass es unter dem Druck der [in die Pfeife geblasenen] Luft das Mundstück genau und gleichmäßig schließt; diese Arbeit erfordert in jedem Fall einen geübten Handwerker. 5) Die Härte des Tons, die [bei herkömmlichen *aufschlagenden Zungenpfeifen*] auf den Anprall des Plättchens auf Metall zurückzuführen ist, wird weicher, wenn der Rand des Mundstücks mit Stanniol, Papier, Leder oder Pergament bedeckt oder wenn die Kehle selbst aus Zinn, Elfenbein oder ölprägniertem Holz gemacht wird. 6) Die Elastizität eines metallenen Plättchens wird vermindert, indem man es mehr oder weniger stark erhitzt oder durch eine Zunge aus weichem Kupfer ersetzt. Eine elfenbeinerne oder aus dem Kieferknochen eines Wals hergestellte und eingesetzte Zunge würde ohne Zweifel einen Ton erzeugen, der der menschlichen Stimme näher käme, doch dass solche [Materialien] – wenigstens bei Kirchenorgeln – sinnvoll eingesetzt werden können, verhindert deren zu große Veränderung durch Luftfeuchtigkeit und Temperaturschwankungen.«

## § 25.

»Um jene Härte bzw. Rauheit des Tons durch den heftigen Anprall des Mundstückventils [des *Zungenblatts*] und dessen durch Grünspan resultierenden Ungenauigkeit vollständig zu vermeiden, und um ferner den künstlichen Artikulationsapparat dem natürlichen stärker analog zu konstruieren, habe ich den Rand des Mundstücks so ausgestattet [vgl. Abb. 7.2, Fig. 6–8], dass der breiteste, äußerste Teil des Plättchens im Zwischenraum am Rand des Mundstücks sich ganz genau passend, doch ohne Reibung, auf und ab bewegen bzw. schwingen kann. Ein gewisser Teil desselben freilich, unterhalb des Stiftes, der die Tonhöhe regelt [die *Stimmkrücke*] liegt in derselben Weise auf dem Rand des Mundstücks wie bei der herkömmlichen Orgel [wie bei der herkömmlichen *aufschlagenden Zungenpfeife*]. Auf diese Weise lässt sich der Ton nicht nur regeln, sondern ver-

liert auch seine Rauheit und gleicht sich der menschlichen Stimme viel stärker an. Auch der Artukulationsapparat selbst wird auf diese Weise dem natürlichen so weit angeglichen, dass jener [künstliche] genau so wie dieser [natürliche] durch die Luftröhre angeblasen, dann, wenn die Luft aus dem Halse zwischen dem Mundstück und dem Plättchen entweicht [wenn der Luftstrom durch den Spalt zwischen *Zungenblatt* und *Kehle* in den *Resonator* eintritt], einem dem menschlichen Laut täuschend ähnlichen hervorbringen kann, sofern das Plättchen [immer wieder] in die Lage zurückkehrt, die Abb. 1 zeigt. Die Schwierigkeit dieser Konstruktion gegenüber der herkömmlichen besteht darin, dass derjenige Teil des Plättchens, der das Mundstück überdeckt, diesem Rand exakt entsprechen muss [zwischen *Zungenblatt* und Rand der *Kehle* darf möglichst kein Raum sein], wenn wir einen schönen Ton hören wollen. Um der Stabilität willen muss das andere Ende des Plättchens mit der Kehle fest verlötet werden; damit aber bei diesem Vorgang das andere Ende sich nicht aus der festgesetzten Lage bewegt, pflege ich aus Rauschgold geschnittene Streifen dazwischen zu legen und erhalte so eine exakte Einpassung ohne Reibung. Je breiter im Übrigen das Plättchen und je elastischer es ist, ein desto lauterer Ton wird daraus resultieren. Dieses neue Instrument bringt ebenso wie das herkömmliche ohne aufgesetzten Becher [wohl: *Resonatoren*] den Diphthong ae hervor [sic!], allerdings auf angenehmere Weise als das herkömmliche. Wenn man ihm aber verschiedene Röhren [*Resonatoren*] aufsetzt, werden von demselben Instrument die Vokale a, e, o, u hervorgebracht.«

## § 31.

»Vorausgesetzt, dass ein Organist geübt in seinem Handwerk ist, dürfte es nicht schwer sein, dass er mit der [neu entwickelten] Vox humana der Orgel die Vokale a, e, o ausspricht. Es soll daher gezeigt werden, wie man mit diesen Pfeifen eine hervorragende und ungewöhnliche Wirkung erzeugen kann, wenn beispielsweise der Ton c den Vokal a erzeugt, der Ton d den Vokal e, der Ton e den Vokal o, der Ton fis wieder den Vokal a und so weiter. Wenn die Melodie im Sopran mit schlichten Tönen gespielt und von einem Bass oder aber von einer lieblichen Blockflöte oder Querflöte oder einer Viola da Gamba begleitet wird, so wird ohne Zweifel der Gesang gleich dem einer

menschlichen Stimme durch diese Veränderung [Entwicklung] hörbar. Der Versuch zeigt außerdem, dass wir beim Sprechen, das aus der Ferne kommt, Konsonanten schlechter, Vokalen jedoch besser verstehen, und dies aus dem Grund, dass, wenn diese [Konsonanten und Vokale] abwechselnd aufeinander folgen, wir die Konsonanten ratend einfügen, weil ohne Zweifel die Veränderung der klingenden Luft bei Konsonanten leichter ausbleibt als bei Vokalen. Sogar bei der Wiederholung von etwas Gesprochenem, das wie ein Echo sein soll, ist es besser, man wiederholt die Vokale als die Konsonanten. In der Musik aber, die ich [oben] vorgestellt habe, soll man im Gesang [in der Melodielinie] diejenigen Quinten, Terzen etc. [und anderen Intervalle] vermeiden, die jeweils einen anderen Vokal hervorbringen, um die unangenehmen Diphthonge ae, oe und ao zu vermeiden.«

## Glossar

Im Folgenden werden sämtliche *orgelbaulichen Fachbegriffe*, die in dieser Arbeit Verwendung fanden, knapp erläutert.

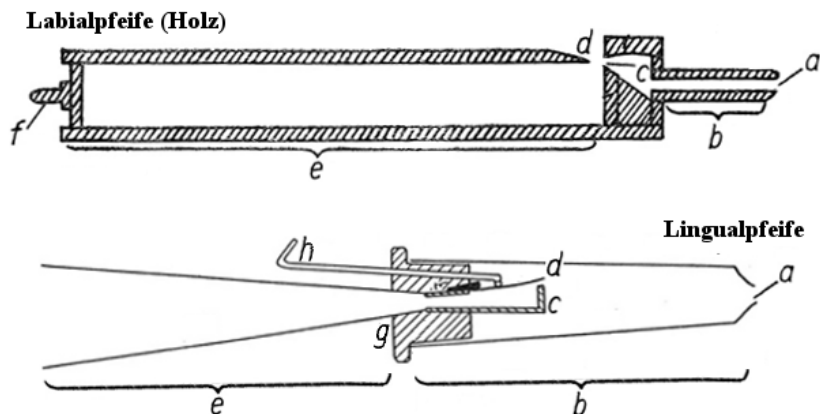
- 8'                   Lies: Acht Fuß. Mit dem historischen Fußmaß wird im Orgelbau die Tonhöhe eines bestimmten  $\rightarrow$  *Registers* bezeichnet, wobei man sich an der jeweils tiefsten (= längsten) Pfeife des Registers orientiert. 8' entsprechen etwa 2,40 m. Register der Achtfußlage klingen in der sog. *Aequallage*: Wird eine Taste  $c^1$  gedrückt, erklingt also auch der Ton  $c^1$ , wohingegen bei einem 4'-Register auf derselben Taste der Ton  $c^2$  erklingt, bei einem 16' folglich  $c^0$ .
- Aufschnitt       bezeichnet die Höhe der *Labiumsöffnung* einer  $\rightarrow$  *Labialpfeife* in Relation zu deren Breite.  $\frac{1}{2}$  bezeichnet also einen Aufschnitt, der doppelt so breit wie hoch ist. Für Orgelpfeifen üblich sind Aufschnittshöhen zwischen  $\frac{1}{5}$  und  $\frac{1}{3}$ .
- Bourdon          siehe *Gedeckt*.
- Crescendotritt   siehe *Schwellwerk*.
- Gedackt          siehe *Gedeckt*.
- Gedeckt           $\rightarrow$  *Labialpfeifen* können an ihrem oberen Ende verschlossen ( $\gg$ gedeckt $\ll$ ) werden, hierdurch klingen sie um eine Oktave tiefer, als es ihrer Körperlänge bei offener Mündung entspräche. Eine 4' lange Pfeife klingt gedeckt also als  $\rightarrow$  8'. Charakteristisch hierbei ist, dass durch die Deckung nur die ungeraden Obertöne verstärkt werden, während bei offenen Pfeifen alle Partialtöne gleichermaßen verstärkt werden. (Abb. 8.15 oben)

Alternative Bezeichnungen für ein gedecktes Labial-Register sind *Gedackt*, *Bourdon*, *Bordun* oder *Stopped Diapason*.

Auch die  $\rightarrow$  *Resonatoren* von  $\rightarrow$  *Lingualpfeifen* können gedeckt werden. Hier dient die Deckung allerdings vordringlich dazu, einen gedämpften, verhaltenen Ton zu erzeugen. Diese Deckung ist dann allerdings nie vollständig, ein gewisser Teil der Schallöffnung muss offen bleiben. Der Grad der Deckung wird als Bruch angegeben: Eine  $\frac{4}{5}$ -Deckung meint

also eine zu vier Fünfteln verschlossene Schallöffnung eines beliebig geformten Resonators.

- Harmonium Auch: *Zungenorgel*. Orgelähnliches Instrument mit ausschließlich durchschlagenden Zungenpfeifen (→ *Lingualpfeifen*). Diese besitzen jedoch weder → *Resonatoren* noch → *Kehlen*, sondern lediglich einen Metallrahmen, in dem das *Zungenblatt* schwingt. Man unterscheidet zwischen *Druckwind*- und *Saugwindharmonien*. Erstere funktionieren analog zu allen übrigen Blasinstrumenten (also auch der Orgel) mit hineingepumpter Luft, die sodann die Pfeifen passiert und durch diese wieder ins Freie strömt. Bei Saugwind-Harmonien wird im Gegensatz hierzu Luft abgepumpt, im Inneren des Instruments also ein Unterdruck erzeugt. Die Luft strömt nun von außen durch die Zungenpfeifen ins Innere des Instruments und wird hierbei angeregt.
- Kanzelle Abgeschlossene Abteilung der → *Windlade*. Bei der klassischen Orgel (bis zum Beginn des 19. Jh.) entspricht die Anzahl der K. einer Windlade deren Tonumfang. Alle einer einzelnen Taste zugehörigen Pfeifen der auf der Windlade stehenden → *Register* befinden sich demnach auf derselben K. und erhalten über ein gemeinsames Ventil ihren → *Wind*. Die Einschaltung der einzelnen Register geschieht über rechtwinklig zu den K. verlaufende *Schleifen*.
- Kehle siehe *Lingualpfeifen*.
- Labialpfeifen Auch: *Lippenpfeifen*. Vgl. Abb. 8.15 oben. Orgelpfeifen mit einem der Blockflöte sehr ähnlichen Funktionsprinzip. Sie bestehen im Wesentlichen aus einem hohlen metallischen Zylinder oder einem hölzernen rechteckigen Kanal, dem *Pfeifenkörper* (e), an dessen unterem Ende sich der *Pfeifenfuß* (b) befindet. Beide Bauteile sind miteinander fest verbunden, in ihrem Inneren aber durch den *Kern* bis auf einen schmalen Spalt voneinander getrennt. An der Position dieser *Kernspalte* (c) ist in den Pfeifenkörper eine rechteckige Öffnung, das *Labium*, geschnitten. Wird nun die Luft von unten in den Pfeifenfuß geleitet, muss sie die Kernspalte passieren, um aus dem Fuß entweichen zu können. Hierbei trifft sie auf die Oberkante des Labiums, das *Oberlabium* (d) und gerät in Schwingungen, die sich im Pfeifenkörper



**Abbildung 8.15:** Schematische Schnittzeichnungen einer Labialpfeife (oben, nach Adelung (1982), 43) und einer Lingualpfeife (unten, nach Lottermoser 1936, 15)

fortsetzen und den Ton erzeugen. Bei → *Gedeckten* ist die Pfeifenmündung durch einen Stöpsel (f) verschlossen.

Lingualpfeifen Auch: *Zungenpfeifen*. Vgl. Abb. 8.15 unten. Die üblichen *aufschlagenden L.* besitzen ein der Klarinette ähnliches Funktionsprinzip. Ein dünnes, schmalrechteckiges Metallstück, das *Zungenblatt* (oder *Zunge*) (d), liegt auf der Oberseite eines längs und an der oberen Schmalseite offenen Rohrsegments, der *Kehle* (c), auf und wird dort durch Luftströmungen in Schwingungen versetzt. Die Kehle steckt mit ihrer offenen oberen Schmalseite in dem sog. *Kopf* (g), einem aus Blei oder Holz bestehenden Block mit mittiger Bohrung, an dessen Oberseite die *Kehle* durch den *Resonator* (e) fortgesetzt wird. Bei *L.* französischer Bauart reicht der *Stiefel* bis über den Ansatz des Resonators hinaus und stützt diesen. Der in diesem Falle im Innern des Stiefels *verschwindende Kopf* wird dann als *Nuss* bezeichnet (vgl. Abb. 6.2, hier reicht der Stiefel bis zum Übergang des konischen zum zylindrischen Teil des Resonators). Der *Kopf* seinerseits steckt im sog. *Stiefel* (b), durch dessen *Fußloch* (a) die Luft strömt, wodurch im Stiefelinneren ein Überdruck entsteht, der das Zungenblatt hochdrückt, damit die Luft durch die *Kehle*



und den *Resonator* ausströmen kann. Durch den *Bernoulli-Effekt* und die Federkraft des *Zungenblatts* wird die Öffnung der *Kehle* sodann wieder verschlossen, der Zyklus beginnt von neuem, die durchströmende Luft wird angeregt, ein Ton entsteht. Um die schwingende Länge des Zungenblattes und damit seine Tonhöhe (je länger, desto tiefer) zu regulieren bedarf es der *Stimmkrücke* (h), einem gebogenen Draht, der verschiebbar von oben auf das obere Ende der Zunge drückt.

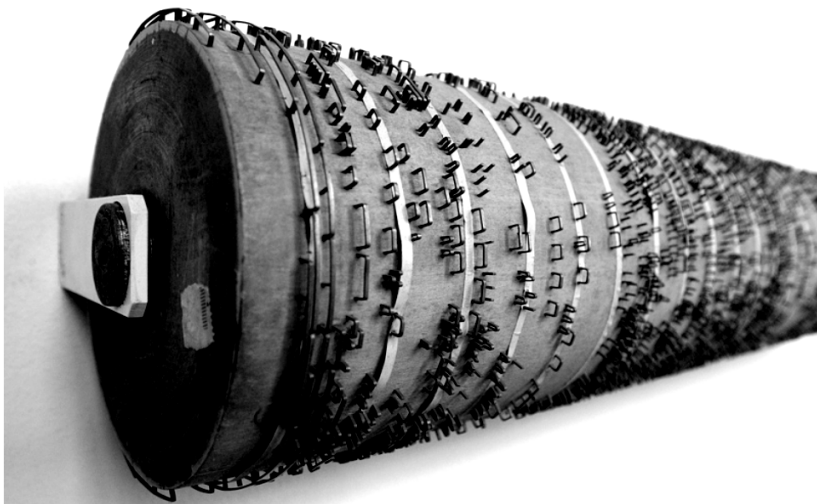
Man unterscheidet grob zwischen sog. *langbecherigen* und *kurzbecherigen* L. Bei ersteren steht die Länge des *Resonators* in direktem Zusammenhang mit dem Grundton der jeweiligen Pfeife und verstärkt diesen. Gleichzeitig werden alle Obertöne gleichmäßig begünstigt. Dies resultiert in einem vollen, Blechblasinstrumenten ähnlichen Klang. *Kurzbecherige* L. zeichnen sich dagegen durch einen schwachen Grundton und große Obertönigkeit aus. Durch die Kürze des *Resonators* werden die Eigengeräusche des *Zungenblattes* nur geringfügig gedämpft, der Klang erscheint schnarrend.

Bei den erheblich selteneren *durchschlagenden Zungen* entspricht die Breite des *Zungenblatts* nicht der der *Kehle*, sondern ist um so viel schmaler, dass sie in der Öffnung der *Kehle* frei schwingen kann und nicht, wie bei *aufschlagenden Lingualpfeifen*, nur immer einen halben Zyklus vollführen kann.

Mensur	Proportion des Durchmessers einer Pfeife zu ihrer klingenden Länge. Diese Proportion ist für den Klang der Pfeifen entscheidend. Bei Labialpfeifen beispielsweise erzeugen weitere Pfeifen einen grundtönigeren, »flötigen«, engere einen obertönigeren, »streichenden« Klang.
Mundorgel	Gruppe von traditionellen Instrumenten aus Asien, erstmals belegt im 9. Jahrhundert v. Chr. M. bestehen aus einer Luftkammer, in die eine Reihe verschieden langer Pfeifen aus Bambus eingesetzt ist. Diese enthalten primitive <i>durchschlagende Zungen</i> (→ <i>Lingualpfeifen</i> ), die sowohl auf Über- als auch auf Unterdruck reagieren. Alle Pfeifen sind mit einem seitlichen Loch versehen. Wird dieses (mit dem Finger) bedeckt, kann die Pfeife klingen. Die <i>Zungen</i> bestehen aus Messing, sind aber im Gegensatz zu westlichen

*durchschlagenden Zungen* mit ihrem Rahmen aus einem Stück gefertigt.

- Positiv      Autonomes → *Werk* einer Orgel, das im Barock mit verhältnismäßig wenigen, eher zarten und leiseren → *Registern* besetzt wurde.
- Quintadena      Aus → *gedeckten Labialpfeifen* verhältnismäßig enger → *Mensur* bestehendes → *Register*.
- Ravalement      Im barocken französischen Orgelbau Erweiterung der Pedalklavatur unterhalb der üblicherweise tiefsten Taste C um bis zu einer Quinte.
- Regal      Spezielle Bauform einer → *Lingualpfeife*, die sich durch kurze, nicht in Proportion zum Grundton der jeweiligen Pfeife stehende → *Resonatoren* auszeichnet. Ergebnis ist ein herber, schnarrender Klang, der durch die wenig gefilterten, zahlreichen Obertöne charakterisiert ist. R. stehen in aller Regel in der → 8'-Lage.
- Register      Reihe von Orgelpfeifen gleicher Klangfarbe, die als Einheit ein- und ausgeschaltet werden kann. In der Regel je Taste/Ton eine Pfeife. Jedes Register ist einem bestimmten → *Werk* zugeordnet, über dessen Klaviatur es angespielt wird.
- Resonator      Auch: *Schallbecher*. Sorgt für eine bessere Abstrahlung der Schallwellen bei → *Lingualpfeifen* und formt durch seine Eigenresonanzen das Klangspektrum der zungeneigenen Schwingung. Zwar kann der Becher theoretisch jede beliebige Länge aufweisen, da die Tonhöhe der Pfeife durch die schwingende Länge des Zungenblatts vorgegeben ist, für volltönende, grundtönige Register sind jedoch Resonatoren notwendig, deren Länge – ähnlich wie die labialen → *Gedekten* –  $\frac{1}{4}$  der Schallwellenlänge des entsprechenden Tones beträgt. Bei sog. *kurzbecherigen* oder → *Regalpfeifen* sind die R. erheblich kürzer und nicht proportional zur Tonhöhe der Pfeife.
- Saugwind      siehe *Harmonium*.
- Schallbecher      siehe *Resonator*.
- Schwellkasten      siehe *Schwellwerk*.



**Abbildung 8.16:** *Stiftwalze einer mechanischen Orgel* (<http://www.deutsches-museum.de/blog/blog-post/2012/10/04/verspielte-maschinen/>) (zuletzt gesichtet am 27.06.2013).

**Schwellwerk** Eine Gruppe von → *Registern* kann innerhalb der Orgel in einem separaten und von allen Seiten verschlossenen Holzkasten stehen, der den Klang stark dämpft. Dessen Vorderseite ist mit einer Art Jalousie versehen, die vom Spieler durch den → *Crescendotritt* (auch: Schwelltritt) geöffnet und geschlossen werden kann, so dass der Klang der betreffenden Pfeifen in seiner Dynamik erheblich beeinflusst werden kann.

**Stiefel** siehe *Lingualpfeifen*.

**Stiftwalze** Steuerung und Programmträger von Spieluhren und mechanischen Orgeln. Hölzerner Zylinder unterschiedlichen Umfanges, auf dessen Oberfläche in definierten Abständen metallene Stifte oder Klammern eingesetzt sind, die während der Drehung der Walze entsprechende Ventile öffnen und so die einzelnen Pfeifen zum Klingen bringen. (Vgl. Abb. 8.16)

**Stimmkrücke** → *Lingualpfeife*

- Temperatur Meint im Orgel- und sonstigen Tasteninstrumentenbau die konkrete Einstimmung des jeweiligen Instrumentes. Anders als beispielsweise bei Saiteninstrumenten ist es bei Tasteninstrumenten nicht möglich, diese vollständig rein zu stimmen, da sich sonst ein Konflikt in der Addition der reinen Oktaven zu derjenigen der reinen Quinten ergibt (»Pythagoreisches Komma«). Als Lösung dieses Widerspruchs wurden seit dem Mittelalter zahlreiche verschiedene T. entwickelt, die sämtlich darauf beruhen, bestimmte harmonische Intervalle absichtlich geringfügig zu verstimmen, um andere möglichst rein zu erhalten. Die moderne *gleichschwebende T.* verstimmt mit Ausnahme der rein erhaltenen Oktave alle Intervalle gleichmäßig, um so ein Spielen in allen Tonarten zu ermöglichen. Bis zum Ende des 18. Jahrhunderts waren verschiedene andere T. gebräuchlich, deren Bestreben es war, nur einige, als entbehrlich erachtete Intervalle zu verstimmen. Bei heutigen Restaurierungen von historischen Orgeln die zwischenzeitlich auch hier angewendete gleichschwebende T. meist wieder revidiert. Da die T.-Systeme der Orgelbauer des Barock beinahe nie verlässlich überliefert sind, muss dies aufgrund von Indizien oder Annahmen geschehen. Hierbei kann es – z. B. durch verschiedene die einzelnen Restaurierungen ausführende Werkstätten – auch bei Orgeln desselben historischen Orgelbauers zu deutlich unterschiedlichen Ergebnissen kommen.
- Tremulant Vorrichtung, mit der der Winddruck (→ *Wind*) einer ganzen Orgel oder eines → *Werks* periodisch variiert werden kann. Hierdurch verändert sich nicht nur die Stärke des jeweils gespielten Tons, sondern auch seine Tonhöhe geringfügig, was zu einem »bebenden« Klang ähnlich dem Vibrato der menschlichen Stimme führt.
- Voce umana Im italienischen Orgelbau der Renaissance und des Barock weit verbreitete → *Register* aus → *Labialpfeifen*. Es handelt sich dabei um je zwei Pfeifen gleicher Tonhöhe pro Taste/Ton, die jedoch geringfügig gegen einander »verstimmt« sind. Der dadurch entstehende, »schwebende« Klang soll an das Tremolo eines singenden Menschen erinnern.
- Vox humana Nach modernem Verständnis bezeichnet VH eine Gruppe von → *Lingualregistern* der → *Regal-Bauweise*. Die genaue

Ausführung insbesondere der  $\rightarrow$  *Resonatoren* unterliegt nur sehr weitgefassten traditionellen Vorgaben; Formen vom einfachen Kegelsegment bis hin zu komplexen zusammengesetzten geometrischen Formen sind belegt. Klangliches Ziel war lange Zeit eine möglichst treffende Nachahmung der (singenden) menschlichen Stimme. Zur Geschichte und Etymologie des Registers siehe Kapitel 2.1.

Werk	Die $\rightarrow$ <i>Register</i> einer Orgel werden in mehrere Gruppen zusammengestellt, die unabhängig voneinander über eigene Klaviaturen angespielt werden können. Die häufigsten Werke sind <i>Hauptwerk</i> , <i>Positiv</i> und <i>Pedal</i> .
Wind	Im Orgelbau die Bezeichnung für (komprimierte) Luft.
Windlade	Herzstück jeder Orgel, das zur Verteilung der Luft ( <i>Wind</i> ) an die einzelnen Pfeifen dient. W. sind hölzerne, recht flache und rechteckige Kästen, auf deren Oberseite die Pfeifen stehen. Ihr Inneres birgt die Luftzufuhr nebst den Ventilen.
Zunge	Kurzform von $\rightarrow$ <i>Zungenblatt</i> .
Zungenblatt	siehe <i>Lingualpfeife</i>
Zungenpfeife	siehe <i>Lingualpfeife</i>



## Audio- und Video-Material

Die hier beschriebenen Klangbeispiele sind unter <http://www.coli.-uni-saarland.de/groups/WB/Phonetics/page.php?id=phonus18> abrufbar. Die *vorangestellte Zahlensignatur* verweist auf den zugehörigen Abschnitt der Arbeit und stellt zugleich den Dateinamen dar.

### 2.1 Canto a tenore

Klangbeispiel einer typischen Canto a tenore-Formation (Video).

### 2.2 Die Vox humana im musikalischen Gebrauch

Zu hören sind drei verschiedene *Voces humanae* (VH):

- Ein Register der französisch-klassischen Bauform in der Orgel der ehem. Abteikirche Marmoutier (Elsass) aus der berühmten Werkstatt des ANDREAS SILBERMANN (1709)
- Ein Register der Gebrüder STUMM (ebenfalls französisch-klassische Bauform) in der Orgel der ev. Schlosskirche Meisenheim/Glan (1767) (auch im Rahmen dieser Dissertation untersucht)<sup>113</sup>
- Das legendäre Register aus der GABLER-Orgel der Basilika in Weingarten (1750)

Vorgestellt werden die Verwendung als solistische Stimme im homophonen Satz (VH als einziges benutztes Register bzw. in Kombination mit dem *Labialregister Gedeckt 8'*) und als Melodie-Register in der klassischen Kombination mit *Gedeckt 8'* und *Tremulant* (die begleitenden Stimmen werden mit anderen Registern gespielt). Ausgewählt wurden Kompositionen, deren Entstehungszeit nahe an der Erbauung der Orgeln liegen.

Bei den verwendeten Klangbeispielen handelt es sich um Ausschnitte aus kommerziellen Audio-CDs. Die detaillierten Quellenangaben finden sich in der im Verzeichnis befindlichen Datei **2.2 README.txt**.

---

<sup>113</sup>Von der im Rahmen der Perzeptionstests besonders gut bewerteten VH in Simmern sind leider keinerlei Aufnahmen verfügbar.

## Verwendete Stimuli

- Kapitel 5:
  - 5.2.2: Experimente mit Labialpfeifen.
    - \* Prinzipalpfeife  $d^2$  mit freiem Labium
    - \* Prinzipalpfeife  $d^2$  mit maximal bedecktem Labium
    - \* Quintadenapfeife  $d^1$  (ohne Deckel =  $d^2$ ) mit freiem Labium
    - \* Quintadenapfeife  $d^1$  (ohne Deckel =  $d^2$ ) mit maximal bedecktem Labium
- Kapitel 6:
  - Perzeptions-Experiment 1: Aufnahmen der 4 verwendeten VH-Register in Originallänge (ca. 5 Sek.), so wie in diesem Experiment verwendet.
    - \* 6.3.1.1-1 Amorbach (VH solo)
    - \* 6.3.1.1-2 Meisenheim (VH solo)
    - \* 6.3.1.1-3 Simmern (VH solo)
    - \* 6.3.1.1-4 Simmern (Cromorne, Trompette)
    - \* 6.3.1.1-5 Waltershausen (VH + Gedeckt)
  - Perzeptions-Experiment 2: Aufnahmen der VH aus Simmern und Waltershausen mit je 400 ms Länge, so wie in diesem Experiment verwendet.
    - \* 6.3.1.2-1 Simmern (VH solo)
    - \* 6.3.1.2-2 Simmern (Cromorne, Trompette)
    - \* 6.3.1.2-3 Waltershausen (VH + Gedeckt)
  - Perzeptions-Experiment 3: Aufnahmen der VH aus Simmern (Kategorie I) sowie der synthetisierten Stimuli der Kategorien II-IIIc, so wie in diesem Experiment verwendet.
    - \* 6.3.1.3-1 Kategorie I: VH-Stimuli aus Simmern
    - \* 6.3.1.3-2 Kategorie II: Klatt-synthetisierte Langokale.  $F_0=100$  Hz,  $F_1-F_3$  nach Simpson (1998).
    - \* 6.3.1.3-3 Kategorie IIIa: Klatt-synthetisierte Stimuli mit  $F_0$  und  $F_1-F_3$  nach Kategorie I.
    - \* 6.3.1.3-4 Kategorie IIIb: Klatt-synthetisierte Stimuli mit  $F_0=100$  Hz und  $F_1-F_2$  nach Kategorie I.



- \* 6.3.1.3-5 Kategorie IIIc: Klatt-synthetisierte Stimuli mit  $F_0$  nach Kategorie I und  $F_1$ - $F_3$  nach Simpson (1998).
- Kapitel 8.3:
  - Perzeptionsexperiment Nr. 2
    - \* 8.3-1 Je eine »Mama«- und eine »Papa«-Realisierung mit den drei unterschiedlichen *Zungenblättern* der Saarbrücker Sprachmaschinen-Replik (6 Stimuli).
    - \* 8.3-2 Je eine »Mama«- und eine »Papa«-Realisierung mit den zwei unterschiedlichen Zungenblättern der Budapester Sprachmaschinen-Replik (4 Stimuli).
  - Perzeptionsexperiment Nr. 3:
    - \* 8.3-3 Alte Windlade: »Mama«-Realisierungen der Saarbrücker Sprachmaschinen-Replik mit zwei unterschiedlichen *Zungenblättern* (Elfenbein breit und Messing schmal). Neun verschiedene Intonationsmuster (18 Stimuli).
    - \* 8.3-4 Neue Windlade (nach historischen Vorbildern): »Mama«-Realisierungen der Saarbrücker Sprachmaschinen-Replik mit zwei unterschiedlichen *Zungenblättern* (Elfenbein breit und Messing schmal). Neun verschiedene Intonationsmuster (18 Stimuli).



Issues of PHONUS published between 1995 and 2015

Barry, W.J., B. Möbius & J. Trouvain (eds.) (2015). PHONUS 18: Fabian Brackhane: »Kann was natürlicher, als Vox humana, klingen?« – Ein Beitrag zur Geschichte der mechanischen Sprachsynthese. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J., B. Möbius & J. Trouvain (eds.) (2014). PHONUS 17: Eva Lasarczyk: *Empirical evaluation of the articulatory synthesizer VocalTractLab as a discovery tool for phonetic research: Articulatory-acoustic investigations of paralinguistic speech phenomena*. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J. & J. Trouvain (eds.) (2011). PHONUS 16. *In memoriam Wolfgang von Kempelen*. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J. (ed.) (2009). PHONUS 15: Manfred Pützer: *Die Rolle kortikaler und subkortikaler Strukturen bei der Initiierung und Produktion differenzierter CV-Silbenwiederholungen. Eine fMRT-Studie*. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J. & J. Trouvain (eds.) (2008). PHONUS 14: Dominika Oliver: *Modeling Polish Intonation for Speech Synthesis*. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J. (ed.) (2008). PHONUS 13: Manfred Pützer: *Stimmqualität und Artikulation bei Dysarthrophonien*. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J. (ed.) (2007). PHONUS 12: Bistra Andreeva: *Zur Phonetik und Phonologie der Intonation der Sofioter-Varietät des Bulgarischen*. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J. & J. Trouvain (eds.) (2007). PHONUS 11: Roland Marti: *ó w dolno-serbsćinje (ó in Lower Sorbian (ó im Niedersorbischen)*. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J. & J. Trouvain (eds.) (2006). PHONUS 10: Caren Brinckmann: *Improving Prosody Prediction for Speech Synthesis – With and Without Symbolic Prosody Features*. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J. & J. Trouvain (eds.) (2005). PHONUS 9. Saarbrücken: Institute of Phonetics, Saarland University.

Barry, W.J. (ed.) (2004). PHONUS 8: Jürgen Trouvain: *Tempo Variation in Speech Production. Implications for Speech Synthesis*. Saarbrücken: Institute of Phonetics, Saarland University of the Saarland.

- Barry, W.J. (ed.) (2004). PHONUS 7: Marc Schröder: *Speech and Emotion Research: An overview of research frameworks and a dimensional approach to emotional speech synthesis*. Saarbrücken: Institute of Phonetics, Saarland University.
- Barry, W.J. & M. Pützer (eds.) (2002). PHONUS 6: *Festschrift für Max Mangold zum 80. Geburtstag*. Saarbrücken: Institute of Phonetics, Saarland University.
- Barry, W.J. & J. Koreman, with K. Kirchhoff (eds.) (2000). PHONUS 5. Saarbrücken: Institute of Phonetics, Saarland University.
- Barry, W.J. & J. Koreman (eds.) (1999). PHONUS 4. Saarbrücken: Institute of Phonetics, Saarland University.
- Barry, W.J. & J. Koreman (eds.) (1997). PHONUS 3. Saarbrücken: Institute of Phonetics, Saarland University.
- Barry, W.J. & A. Addison (eds.) (1996). PHONUS 2. Saarbrücken: Institute of Phonetics, Saarland University.
- Barry, W.J. & J. Koreman (eds.) (1995). PHONUS 1. Saarbrücken: Institute of Phonetics, Saarland University.

*Electronic versions (PDFs) are available online at:*

*<http://www.coli.uni-saarland.de> > Research Groups > Phonetics > Phonus*

**Fabian Brackhane**

30. Mai 1980	geboren in Regensburg
2000	Abitur in Herford (Westfalen)
2000-01	Zivildienst in Ulm
2001-09	Studium der Phonetik, Musikwissenschaft und Informationswissenschaft an der Univer- sität des Saarlandes, Saarbrücken
2009	Magister artium
2015	Promotion